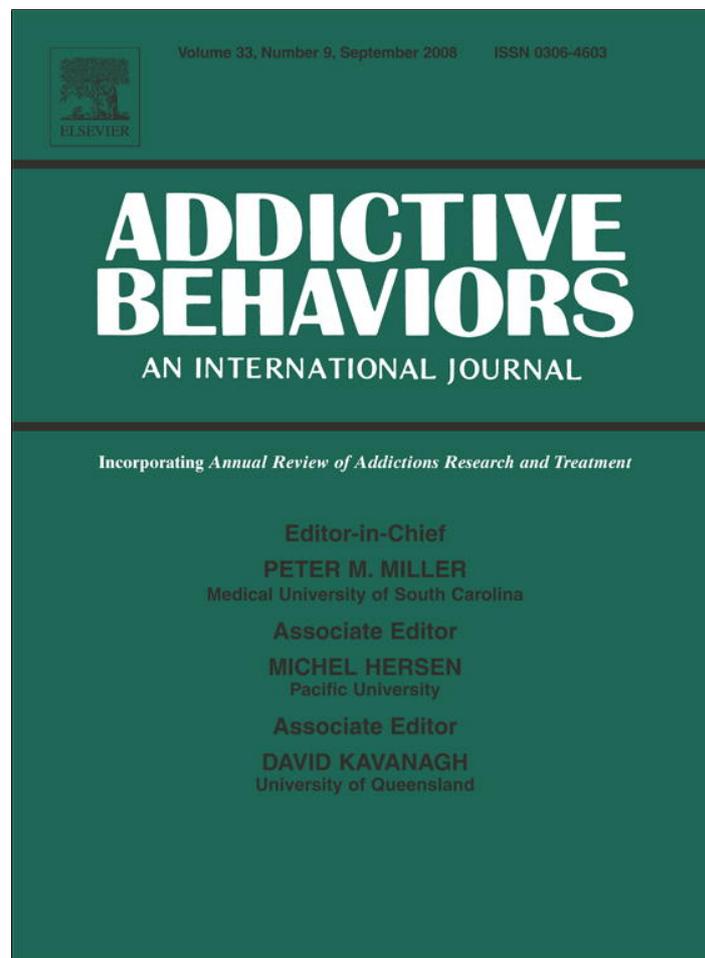


Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>

Contents lists available at [ScienceDirect](#)

Addictive Behaviors



Short communication

A brief Rutgers Alcohol Problem Index with less potential for bias

Mitch Earleywine^{a,*}, Joseph W. LaBrie^b, Eric R. Pedersen^b^a University at Albany, State University of New York, USA^b Loyola Marymount University, Los Angeles, USA

ARTICLE INFO

Keywords:

Alcohol
Problems
Differential item function
Bias
Drinking
Gender differences

ABSTRACT

The Rutgers Alcohol Problem Index (RAPI), a popular measure of alcohol-related problems in adolescents, varies with many theoretically-relevant measures of individual differences, including sex. The sex differences in RAPI scores fit many models of alcohol problems but could also arise from biased items. In addition, a short form could increase the scale's utility. The current study examined RAPI scores, an additional inventory of problem drinking, and measures of alcohol consumption in over 2000 college student drinkers. Analyses revealed items that functioned differentially for men and women. Dropping these items created a shorter scale with almost identical psychometric properties but less potential for bias. Correlations with drinking habits and drinking problems were the same as those for the full scale, and the size of the effect for the difference between men and women's responses remained essentially the same. These results confirm previous work using different analytic approaches, and suggest that a short form of the RAPI could prove helpful in future research. In addition, these data suggest that analyses of differential item functioning in other scales can reveal important information about the measurement of drug problems.

© 2008 Elsevier Ltd. All rights reserved.

1. A brief Rutgers Alcohol Problem Index with less potential for bias

The Rutgers Alcohol Problem Index (RAPI; White & Labouvie, 1989), a common measure of the efficacy of interventions with college students (e.g., Borsari & Carey, 2005; LaBrie, Lamb, Pedersen, & Quinlan, 2006; LaBrie, Pedersen, Lamb, & Quinlan, in press; Neighbors, Larimer, & Lewis, 2004), varies with numerous alcohol-related constructs (Anderson, Simmons, Martens, Ferrier, & Sheely, 2006; Broman, 2005; Danielson, Overholser, & Butt, 2003; LaBrie et al., in press). Alcohol consumption and problems can vary with biological sex. The association between consumption and consequences appears larger for women in some studies (e.g., Murphy, McDevitt-Murphy, & Barnett, 2005; Stewart, Hutson, & Connors, 2006), but not others (Neve, Lemmens, & Drop, 1997; Plant, Miller, Thornton, Plant, & Bloomfield, 2000). The magnitude of the sex differences in these problems is impossible to estimate if items show bias against one sex. Sex-moderated links between predictors of negative consequences and alcohol problems could also arise from item bias. For these reasons, an examination of sex bias in the assessment of alcohol problems seems potentially heuristic. Previous work (Neal, Corbin, & Fromme, 2006) identified two RAPI items with the potential for sex bias using a 2-parameter IRT analysis; we chose alternative techniques that require different assumptions.

One approach involves differential item functioning (DIF), a necessary condition for bias that occurs when an item unfairly favors one group. Any item that is more likely to indicate pathology in one group when both groups are matched on level of pathology functions differentially. For example, the Center for Epidemiologic Studies Depression scale (CES-D; Radloff, 1977) has excellent psychometric properties but an item about crying functions differently across the sexes. Women are more likely to endorse that they had crying spells than men who are equally depressed (Gelin & Zumbo, 2003). We looked for comparable biases on the RAPI.

* Corresponding author. University at Albany, State University of New York, SS369, Department of Psychology, Albany, New York, 12222, USA. Tel.: +1 518 338 7455.

E-mail address: mearleywine@albany.edu (M. Earleywine).

2. Method

2.1. Participants

The 2333 participants all drank at least one drink per month. They were selected from a large sample of college students from a West Coast university who responded to questionnaires as part of a set of studies on alcohol. They included 903 men (39%) and 1430 women (61%), with an average age of 19.1 years ($SD=1.30$). Ethnic affiliations included 1560 (67%) Caucasians, 289 (12%) Latinos, 148 (6%) Asians, 61 (3%) African Americans, and 272 (11%) of mixed race. Three participants did not report ethnicity.

2.2. Measures

2.2.1. Drinking habits

2.2.1.1. Quantity and frequency. Participants reported the average number of standard drinks per occasion and the average number of drinking occasions per month in the past year. They averaged 40.9 drinks per month ($SD=43.1$).

2.2.1.2. Timeline followback. A subset of participants (969) reported drinks consumed each day for the previous month (Sobell & Sobell, 1992), averaging 46.0 ($SD=51.4$).

2.2.2. Drinking problems

2.2.2.1. Rutgers Alcohol Problem Index (RAPI). The 23-item RAPI assesses the frequency of alcohol-related negative consequences (White & Labouvie, 1989), scored from 0 (never) to 4 (10 or more times). Cronbach's alpha was .88. Scores averaged 5.4 ($SD=6.6$).

2.2.2.2. College Alcohol Problem Scale – revised (CAPS-r). A subset of 1095 respondents completed this 8-item scale, which assesses alcohol problems like feeling sad, blue, or depressed from drinking (Maddock, LaForge, Rossi, O'Hare, 2001). Items range from 1 (never/almost never) to 5 (very often). Internal consistency was .81. Scores averaged 11.3 ($SD=4.3$).

3. Analyses

The first two approaches to detecting DIF rely on the idea that men and women with equal alcohol problems should endorse comparable values on an item. If item scores differ dramatically when groups are matched on total scores, the item functions differentially. Popular indices of DIF include the Mantel Chi-square (Mantel, 1963, Zwick, Donoghue, & Grima, 1993; Zwick, Thayer,

Table 1

RAPI items included in full scale (all items) and brief versions (bold items removed)

1	Not able to do your homework or study for a test
2	Got into fights with other people (friends, relatives, strangers)
3	Missed out on other things because you spent too much money on alcohol (MR)
4	Went to work or school high or drunk (WS NI)
5	Caused shame or embarrassment to someone
6	Neglected your responsibilities
7	Relatives avoided you
8	Felt that you needed more alcohol than you used to in order to get the same effect
9	Tried to control your drinking (tried to drink only at certain times of the day or in certain places, that is, tried to change your pattern of drinking)
10	Had withdrawal symptoms, that is, felt sick because you stopped or cut down on drinking
11	Noticed a change in your personality (NI)
12	Felt that you had a problem with alcohol
13	Missed a day (or part of a day) of school or work
14	Wanted to stop drinking but couldn't
15	Suddenly found yourself in a place that you could not remember getting to
16	Passed out or fainted suddenly (WR)
17	Had a fight, argument or bad feeling with a friend (MS)
18	Had a fight, argument or bad feeling with a family member (WR)
19	Kept drinking when you promised yourself not to (MS)
20	Felt you were going crazy
21	Had a bad time
22	Felt physically or psychologically dependent on alcohol
23	Was told by a friend, neighbor or relative to stop or cut down drinking

WS=item functioned differentially against women in the Standardized Liu–Agresti Cumulative Common Log-Odds Ratio and Mantel analyses.

MS=item functioned differentially against women in the Standardized Liu–Agresti Cumulative Common Log-Odds Ratio and Mantel analyses.

NI=item failed to pass the interval-scaling assumption of the Rasch model.

WR=item functioned differentially against women in the Rasch difficulty estimates.

MR=item functioned differentially against men in the Rasch difficulty estimates.

Table 2

Correlations among RAPI, short forms, drinking habits, and drinking problems

1. RAPI	2	3	4	5
2. S-RAPI	98			
3. TLFB	40	41		
4. QF	45	46	73	
5. CAPS-r	64	63	38	41

RAPI = total Rutgers Alcohol Problem Inventory.

S-RAPI = 16-item short form based on any DIF procedure.

TLFB = Timeline Followback drinks in previous month ($N=969$).

QF = Average drinks per month based on reported quantity and frequency.

CAPS-r = College Alcohol Problems Scale – revised ($N=1095$).

Decimal points omitted.

All p -values less than .001.

& Mazzeo, 1997) and Standardized Liu–Agresti Log-Odds ratios Ratios (SLACCLOR; Liu & Agresti, 1996; Penfield & Algina, 2003). An alternative approach (Rasch modeling) ranks items according to their ease of endorsement along a latent construct of alcohol problems. DIF appears when groups differ in the difficulty of endorsement of an item. For example, previous work shows that women endorse the alcohol expectancy item “sweet alcoholic drinks taste good” at lower levels of a latent construct of positive alcohol expectancies than men (Mackintosh, Earlywine, & Dunn, 2006). Only statistics significant at $p < .001$ served as signs of DIF.

4. Results

4.1. Mantel Chi-square and Standardized Liu–Agresti Cumulative Common Log-Odds Ratio (SLACCLOR)

Mantel Chi-square and Standardized Liu–Agresti Cumulative Common Log-Odds Ratio (SLACCLOR) both identified the same three items. Item 4 (went to work or school high or drunk) appeared biased against women, who gave higher answers than men with comparable total scores (MCS=11.15; SLACCLOR=3.38, all $ps < .001$). Item 17 (had a fight, argument, or bad feelings with a friend; MCS=10.68; SLACCLOR=-3.31) and Item 19 (kept drinking when you promised yourself not to; MCS=17.15; SLACCLOR=-4.41) appeared biased against men (all $ps < .001$). See Table 1.

4.2. Rasch difficulty estimates

We removed items that violated the assumption of unidimensional, interval scaling, and then examined DIF based on item difficulty (Bond & Fox, 2001; Linacre, 2004; Smith & Miao, 1994). Items 4 (went to work or school high or drunk), 11 (noticed a change in your personality) and 19 (kept drinking when you promised yourself not to) did not satisfy requirements for interval scaling (Linacre, 2004). An analysis of the remaining items showed DIF on 3, 16, and 18. Men endorsed item 3 (missed out on things because you spent too much money on alcohol) at a lower level of problems. (The same problem appeared in high school students (Neal et al., 2006)). Women endorsed item 16 (passed out or fainted suddenly) and 18 (had a fight, argument, or bad feelings with a family member) at a lower level of problems ($t=-3.22, 3.53, \text{ and } 4.41$, respectively, all $ps < .001$). See Table 1.

4.2.1. Revised scale

Removing items that did not fit the Rasch model or that showed bias left 16 items for the short RAPI (S-RAPI). Comparisons of correlated correlation coefficients (Meng, Rosenthal & Rubin, 1992) revealed that the sum of the dropped items correlated significantly less than the S-RAPI with the CAPS (.56 vs. .63; $Z=7.50, p < .001$), TLFB (.31 vs. .41; $Z=4.92, p < .001$) and the Quantity-Frequency measure (.36 vs. .46; $Z=4.30, p < .001$). Internal consistency dropped from .88 to .85. The S-RAPI had correlations with all other measures that were within .01 of the correlations with the full scale, despite a 30% reduction in length (see Table 2.). The sexes differed on the RAPI and S-RAPI ($t(2,331)=8.64 \text{ and } 8.67$, respectively, $p < .001$). Effect sizes were $d = .43$ for both scales (Table 3).

Table 3

Means and standard deviations for men and women on the RAPI and short form

	Sex	Mean	SD
*RAPI	Men	6.9	7.9
	Women	4.5	5.5
*S-RAPI	Men	5.1	5.8
	Women	3.3	4.2

 N for Men=903. N for Women=1430.

S-RAPI=brief scale based on DIF analyses with Standardized Liu–Agresti Cumulative Common Log-Odds Ratio and Mantel analyses or Rasch analyses.

* Sex differences significant at $p < .001$.

5. Discussion

Multiple techniques revealed potential sex biases in the Rutgers Alcohol Problem Index (RAPI). Removing biased items did not alter internal consistency of the scale, its correlates with drinking habits and problems, or the size of sex differences in problems. The content of the biased items suggests that sex roles might contribute to indices of alcohol problems. Three items were potentially biased against women: item 4 (went to work or school high or drunk), item 16 (passed out or fainted suddenly) and 18 (had a fight, argument, or bad feelings with a family member). Item 4's results may stem from different definitions of 'high' or 'drunk'. Men might drink before obligations but not consider themselves high or drunk, and women require fewer drinks to get drunk (Plant et al., 2000). Item 16 (passed out or fainted suddenly) showed a comparable bias that might arise from lower body mass or dieting. A lower body mass could lead to a higher blood alcohol concentration, increasing the chance of passing out. The higher rates of chronic dieting in women (Cachelin & Regan, 2006) might also contribute. Item 18's bias (had a fight, argument, or bad feelings with a family member) may suggest that relatives are more likely to mention problem drinking to women than to men. Women might also be more sensitive to relational consequences (Gleason, 1994; Vince-Whitman & Cretella, 1999).

Three items showed bias against men. Item 3 (missed out on things because you spent too much money on alcohol) may stem from men drinking larger quantities or purchasing drinks for women. Item 17 (had a fight, argument, or bad feelings with a friend) might rest on what qualifies as a fight, argument, or bad feeling to women that might not qualify to men. Item 19 (kept drinking when you promised yourself not to) suggests that men are less likely to promise to decrease drinking until they have reached a higher level of problems in the first place.

The clarion call for further research and replication is particularly relevant here. Few studies of differential item function have been replicated. The RAPI is used widely, so a replication of these DIF analyses with other data sets should prove relatively straightforward. In addition, the range of ages and education on these participants is limited. The RAPI was essentially designed for this population, but assuming that DIF would be comparable in participants of markedly different ages or educational status is probably unreasonable. Another study of the RAPI found DIF for an item that was also problematic in the data from this study, "missed out on things because you spent too much money on alcohol". In contrast, though, the item "had a bad time" showed gender-related DIF in that study but not in the current one (Neal et al., 2006).

This new, brief version of the RAPI has a main advantage involving less potential for sex bias, allowing more confidence in estimates of sex differences in problems. This shorter version might also prove easier for participants to complete, particularly as part of a larger survey. A briefer scale might encourage researchers who do not normally assess drinking problems in college students to expand their work into this area. In addition, these results have implications for other forms of bias with the RAPI and other measures of drug-related problems. DIF and bias can appear across sexes, but also across ethnic groups. Although the current sample lacked adequate power to examine DIF across ethnic groups, such work could prove particularly helpful. Previous studies have shown DIF across ethnic groups for items assessing various forms of psychopathology (Cauffman & MacIntosh, 2006), and the potential for DIF across ethnic groups for the RAPI and other measures of drug problems seems worthy of examination. The presence of sex-related DIF in the RAPI also suggests that other measures of drug problems might show sex-related DIF. As investigations of DIF progress, the field can improve estimates of group differences on measures of drug problems with the assurance that such differences stem from genuine deviations in true scores rather than unique aspects of individual items.

Acknowledgments

This research was funded by a grant from the Alcoholic Beverage Medical Research Foundation and Grant U18 AA015451-01 from the National Institute of Alcohol Abuse and Alcoholism. Special thanks to Elana B. Gordis for continued support.

References

- Anderson, D., Simmons, A. M., Martens, M. P., Ferrier, A. G., & Sheehy, M. J. (2006). The relationship between disordered eating behavior and drinking motives in college-age women. *Eating Behaviors*, 7, 419–422.
- Bond, T. G., & Fox, C. M. (2001). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Borsari, B., & Carey, K. (2005). Two brief alcohol interventions for mandated college students. *Psychology of Addictive Behaviors*, 19, 296–302.
- Broman, C. L. (2005). Stress, race, and substance use in college. *College Student Journal*, 39, 340–352.
- Cachelin, F. M., & Regan, P. C. (2006). Prevalence and correlates of chronic dieting in a multi-ethnic U.S. community sample. *Eating and Weight Disorders*, 11, 91–99.
- Cauffman, E., & MacIntosh, R. (2006). A Rasch differential item functioning analysis of the Massachusetts Youth Screening Instrument: Identifying race and sex differential item functioning among juvenile offenders. *Educational and Psychological Measurement*, 66, 502–521.
- Danielson, C. K., Overholser, J. C., & Butt, Z. (2003). Association of substance abuse and depression among adolescent psychiatric inpatients. *Canadian Journal of Psychiatry*, 48, 762–765.
- Gelin, M. N., & Zumbo, B. D. (2003). Differential item functioning results may change depending on how an item is scored: An illustration with the center for epidemiologic studies depression scale. *Educational and Psychological Measurement*, 63, 65–74.
- Gleason, N. A. (1994). Preventing alcohol abuse by college women: A relational perspective. *Journal of American College Student Health*, 43, 15–25.
- LaBrie, J. W., Lamb, T. F., Pedersen, E. R., & Quinlan, T. (2006). A group Motivational Interviewing intervention reduces drinking and alcohol-related consequences in adjudicated college students. *Journal of College Student Development*, 47, 267–280.
- LaBrie, J. W., Pedersen, E. R., Lamb, T., & Quinlan, T., in press A campus-based group motivational intervention reduces problematic drinking in freshmen male college students. *Addictive Behaviors*, 32, 889–901.
- Linacre, J. M. (2004). *WINSTEPS: Rasch Measurement Computer 2007 Program*. Chicago: Winsteps.com.
- Liu, I.-M., & Agresti, A. (1996). Mantel–Haenszel-type inference for cumulative odds ratios with a stratified ordinal response. *Biometrics*, 52, 1223–1234.
- Mackintosh, M. A., Earleywine, M., & Dunn, M. E. (2006). Alcohol expectancies for social facilitation: A short form with decreased bias. *Addictive Behaviors*, 31, 1536–1546.

- Maddock, J. E., LaForge, R. G., Rossi, J. S., & O'Hare, T. (2001). The college alcohol problems scale. *Addictive Behaviors*, 26, 385–398.
- Mantel, N. (1963). Chi-square tests with one degree of freedom: Extension of the Mantel–Haenszel procedure. *Journal of the American Statistical Association*, 58, 690–700.
- Meng, X. L., Rosenthal, R., & Rubin, D. (1992). Comparing correlated correlation coefficients. *Psychological Bulletin*, 111, 172–175.
- Murphy, J. G., McDevitt-Murphy, M. E., & Barnett, N. P. (2005). Drink and be merry? Sex, life satisfaction, and alcohol consumption among college students. *Psychology of Addictive Behaviors*, 19, 184–191.
- 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.
- Neighbors, C., Larimer, M. E., & Lewis, M. A. (2004). Targeting misperceptions of descriptive drinking norms: Efficacy of a computer-delivered personalized normative feedback intervention. *Journal of Consulting and Clinical Psychology*, 72, 434–447.
- Neve, R. J., Lemmens, P. H., & Drop, M. J. (1997). Sex differences in alcohol use and alcohol problems: Mediation by social roles and sex-role attitudes. *Substance Use and Misuse*, 32, 1439–1459.
- Penfield, R. D., & Algina, J. (2003). Applying the Liu–Agresti estimator of the cumulative common odds ratio to DIF detection in polytomous items. *Journal of Educational Measurement*, 40, 353–370.
- Plant, M., Miller, P., Thornton, C., Plant, M., & Bloomfield, K. (2000). Life stage, alcohol consumption patterns, alcohol-related consequences, and sex. *Substance Abuse*, 21, 265–281.
- Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, 3, 385–401.
- Smith, R. M., & Miao, C. Y. (1994). Assessing unidimensionality for Rasch measurement. In M. Wilson (Ed.), *Objective measurement: Theory into practice*, vol. 2. Norwood, NJ: Ablex.
- Sobell, L. C., & Sobell, M. B. (1992). Timeline followback: A technique for assessing self-reported alcohol consumption. In R. Z. Litten, & J. P. Allen (Eds.), *Measuring alcohol consumption: Psychosocial and biological methods* (pp. 41–72). Totowa, NJ: Humana Press.
- Stewart, S. H., Hutson, A., & Connors, G. J. (2006). Exploration of the relationship between drinking intensity and quality of life. *American Journal of the Addictions*, 15, 356–361.
- Vince-Whitman, C., & Cretella, M. (1999). Alcohol use by college women: Patterns, reasons, results, and prevention. *Catalyst: A Publication of the Higher Education Center for Alcohol and Other Drug Prevention*, 5, 4–5.
- White, H. R., & Labouvie, E. W. (1989). Towards the assessment of adolescent problem drinking. *Journal of Studies on Alcohol*, 50, 30–37.
- Zwick, R., Donoghue, J. R., & Grima, A. (1993). Assessment of differential item functioning for performance tasks. *Journal of Educational Measurement*, 30, 233–251.
- Zwick, R., Thayer, D. T., & Mazzeo, J. (1997). Descriptive and inferential procedures for assessing differential item functioning in polytomous items. *Applied Measurement in Education*, 10, 321–334.