Addictive Behaviors Reports 8 (2018) 71–78

Contents lists available at ScienceDirect

Addictive Behaviors Reports

journal homepage: www.elsevier.com/locate/abrep

A psychometric evaluation of the Protective Behavioral Strategies Scale-20 among internet samples of adult drinkers

Dylan K. Richards*, Reyna P. Puentes, Rubi Gonzales, Juliana Cardoso Smith, Craig A. Field, Osvaldo F. Morera

Latino Alcohol and Health Disparities Research (LAHDR) Center, Department of Psychology, University of Texas at El Paso, 500 W. University Ave., El Paso, TX 79968, United States of America

ARTICLE INFO

Keywords:
Alcohol
Adults
Protective behavioral strategies
Psychometrics
Mechanical Turk

ABSTRACT

Protective behavioral strategies (PBS) are most commonly defined as behaviors that are used while drinking to reduce alcohol use and/or limit alcohol-related problems. Few studies have examined and quantified PBS use among non-college student populations. The purpose of the present two studies was to evaluate the psychometric properties of the Protective Behavioral Strategies Scale-20 (PBSS-20; Treloar, Martens, & McCarthy, 2015) among internet samples of adult drinkers. In the first study, we conducted an exploratory factor analysis of the PBSS-20 with a sample (n = 360) of adult drinkers who were recruited from Mechanical Turk. We then conducted a second study that recruited adult drinkers from Mechanical Turk and randomly split the data in half. With the first split-half sample (n = 339), we conducted a confirmatory factor analysis of the PBSS-20 and assessed the internal consistency and concurrent validity of the subscales. With the second split-half sample (n = 338), we tested measurement invariance across gender. The results support a three-factor structure of the PBSS-20 that is similar to what has been found among college students. However, six items were dropped and two Serious Harm Reduction items loaded best onto the Manner of Drinking factor. Furthermore, two subscales demonstrated adequate internal consistency and all three subscale were negatively associated with alcohol-related outcomes. Similar to college students, there was lack of measurement invariance across gender. We discuss the implications of the present findings in extending research on PBS to the more general population of U.S. adult drinkers.

1. Introduction

Protective behavioral strategies (PBS) are most commonly defined as behaviors that are used *while* drinking to reduce alcohol use and/or limit alcohol-related problems (Martens et al., 2005). Notably, others (e.g., Novik & Boekeloo, 2011; Sugarman & Carey, 2007) have defined PBS more broadly to include alcohol-avoidance behaviors; however, relatively few studies (~21% of PBS literature) have used this broader definition and there are benefits to both the conceptualization and measurement of PBS by using the narrower definition (Pearson, 2013). In a review of the literature on PBS use among college students, Pearson (2013) notes the increasing evidence for the cross-sectional relationships between more frequent use of PBS and less alcohol use and fewer alcohol-related problems. Furthermore, there is some evidence for PBS as proximal outcomes (i.e., mediators) of interventions that target the reduction of alcohol-related problems among college students (Barnett, Murphy, Colby, & Monti, 2007; Larimer et al., 2007; Murphy et al., 2012). Although there is no reason to suspect that PBS are not used by or may not be effective among non-college student populations, studies on PBS have almost exclusively been conducted with college student samples (Pearson, 2013). Thus, research examining PBS use among other populations of drinkers is warranted, as PBS may be useful in reducing the public health burden of alcohol misuse and related problems among these populations. However, to extend research on PBS use to non-college student populations, reliable and valid measures of PBS use among these populations are needed.

It may be that existing measures of PBS use that were developed and validated among college students also validly assess PBS use among other populations. The Protective Behavioral Strategies Scale (PBSS; Martens et al., 2005) is the most widely used and well-validated measure of PBS use among college students (Prince, Carey, & Maisto, 2013). Previous research among college students (e.g., Martens et al., 2005; Martens, Pederson, LaBrie, Ferrier, & Cimini, 2007) supports a three-factor model of the PBSS that includes the following subscales:

* Corresponding author at: Department of Psychology, University of Texas at El Paso, 500 W. University Ave., El Paso, TX 79968, United States of America.
E-mail address: Dkrichards2@miners.utep.edu (D.K. Richards).

https://doi.org/10.1016/j.abrep.2018.08.002
Received 8 February 2018; Received in revised form 17 July 2018; Accepted 2 August 2018
Available online 03 August 2018
2352-8532/ © 2018 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/BY-NC-ND/4.0/).

Table 1
Descriptive statistics for the socio-demographic and alcohol-related variables by sample.

Categorical variables	Sample 1 (n = 360)		Sample 2 (n = 339)		Sample 3 (n = 338)		
	n	%	*n*	%	*n*	%	
Gender							$\chi^2 = 4.59$, $df = 2$, $p > .05$
Female	203	56.4	184	54.3	210	62.1	
Male	157	43.6	155	45.7	128	37.9	
Marital status							$\chi^2 = 6.84$, $df = 8$, $p > .05$
Single (never married)	118	32.8	99	29.2	118	34.9	
Married	153	42.5	164	48.4	140	41.4	
Divorced	29	8.1	24	7.1	30	8.9	
Separated	6	1.7	4	1.2	8	2.4	
Living with someone	54	15.0	48	14.2	42	12.4	
Ethnicity							$\chi^2 = 8.97$, $df = 10$, $p > .05$
Hispanic	28	7.8	22	6.5	21	6.2	
White	294	81.7	272	80.2	280	82.8	
Asian American	15	4.2	16	4.7	20	5.9	
African American	18	5.0	17	5.0	13	3.8	
Native American	1	0.3	2	0.6	0	0.0	
Other	4	1.1	10	2.9	4	1.2	
Income							$\chi^2 = 5.78$, $df = 6$, $p > .05$
< $15,000	37	10.3	23	6.8	24	7.1	
Between $15,000 and $30,000	60	16.7	63	18.6	63	18.6	
Between $30,000 and $50,000	98	27.2	84	24.8	98	29.0	
> $50,000	165	45.8	169	49.9	153	45.3	
Frequency of alcohol use							$\chi^2 = 9.90$, $df = 10$, $p > .05$
About once a month	41	11.4	41	12.1	36	10.7	
Two to three times a month	73	20.3	64	18.9	83	24.6	
Once or twice a week	99	27.5	87	25.7	98	29.0	
Three to four times a week	95	26.4	84	24.8	82	24.3	
Nearly every day	38	10.6	45	13.3	28	8.3	
Once a day or more	14	3.9	18	5.3	11	3.3	

Continuous variables	Sample 1 (n = 360)		Sample 2 (n = 339)		Sample 3 (n = 338)		
	M	*SD*	*M*	*SD*	*M*	*SD*	
Age (years)	37.71	10.70	37.09	9.63	36.68	9.47	$F(2, 1034) = 0.95$, $p > .05$
Typical quantity	3.73	3.56	3.90	4.11	3.47	3.46	$F(2, 1034) = 1.16$, $p > .05$
Heaviest quantity	5.51	4.77	5.71	4.99	5.25	4.28	$F(2, 1034) = 0.79$, $p > .05$
SIP +6 score	5.07	8.33	4.68	7.23	5.08	7.54	$F(2, 1034) = 0.30$, $p > .05$

Note. Typical quantity = number of alcoholic drinks consumed on a typical drinking occasion in the past three months; heaviest quantity = number of alcoholic drinks consumed on the heaviest drinking occasion in the past three months; SIP +6 = Short Inventory of Problems +6.

Stopping/Limiting Drinking (7 items; e.g., "Alternate alcoholic and nonalcoholic drinks"), Manner of Drinking (5 items; e.g., "Avoid drinking games"), and Serious Harm Reduction (3 items; e.g., "Use a designated driver"). Although the three-factor model has been largely supported, studies have found evidence for some lack of fit of the model (e.g., Martens et al., 2007). The Stopping/Limiting Drinking and Manner of Drinking subscales have demonstrated acceptable reliability and all three subscales negatively correlate with alcohol use and alcohol-related problems (e.g., Martens et al., 2005, 2007). The PBSS was recently revised (PBSS-20; Treloar, Martens, & McCarthy, 2015) to improve the content validity of the Serious Harm Reduction subscale, which was expanded from 3 items to 8 items.

Given that the PBSS is the most widely used and well-validated measure of PBS use among college students, it seems to be the best candidate to evaluate as a measure of PBS use among non-college student populations. To our knowledge, only one study (Cadigan, Weaver, McAfee, Herring, & Martens, 2015) examined the psychometric properties of the PBSS among a non-college student sample. In a mostly male (94%) sample of military veterans of the wars in Afghanistan and Iraq, Cadigan et al. (2015) found some support for the three-factor model of the PBSS demonstrated among college students. The authors also found support for the reliability and validity of the Stopping/Limiting Drinking and Manner of Drinking subscales. However, the internal consistency estimate for the Serious Harm Reduction subscale was low ($\alpha = 0.52$) and the subscale did not predict drinking outcomes in hierarchical regressions that controlled for gender (Cadigan et al., 2015).

Despite the importance of the Cadigan et al. (2015) study, further psychometric evaluation of the PBSS among non-college student samples is warranted for several reasons. First, the sample in the Cadigan et al. study was predominantly male and psychometric evaluations of the PBSS should be extended to non-college student samples that are more representative of females. Also, previous research suggests lack of measurement invariance of the PBSS and PBSS-20 across gender (Treloar et al., 2015; Treloar, Martens, & McCarthy, 2014) which Cadigan et al. were unable to test given the low number of females in their sample. Second, the sample in the Cadigan et al. study was a sample of military veterans. Military veterans are more likely to be at-risk drinkers than the general adult population (e.g., Hawkins, Lapham, Kivlahan, & Bradley, 2010) and thus military veterans' PBS use may be inherently different than the general adult population. Third, Cadigan et al. (2015) used the PBSS and improvements have since been made to the Serious Harm Reduction subscale in the PBSS-20. Fourth, Cadigan et al. found some support for the confirmation of the three-factor model of the PBSS among military veterans, but the authors did not explore whether other factor solutions may have provided a better fit to the data. Finally, to our knowledge, no other study has attempted to replicate the factor structure and other psychometric properties of the PBSS-20 with any sample, including college student drinkers.

1.1. The present studies

The aim of the present studies was to evaluate the psychometric properties of the PBSS-20 among internet samples of adult drinkers that were recruited from Amazon's Mechanical Turk (MTurk). MTurk samples have been found to be much more diverse than typical convenience samples of college students (Buhrmester, Kwang, & Gosling, 2011). Samples recruited from MTurk may be useful in attempting to generalize research on the measurement of PBS use to a more general population of adult drinkers in the U.S. In one study, we explored the factor structure of the PBSS-20. In a second study, we randomly split the data in half and, with the first split-half sample, we attempted to confirm the factor structure of the PBSS-20 that was supported in the first study as well as examine the internal consistency and concurrent validity of the subscales. We used the second split-half sample from the second study to test the measurement invariance of the PBSS-20 across gender.

2. Method

2.1. Participants and procedure

In two separate studies, 360 and 677 adult drinkers, respectively, were recruited from MTurk. Data from the second study was randomly split in half: the first split-half sample (i.e., Sample 2; $n = 339$) and the second split-half sample (i.e., Sample 3; $n = 338$). Table 1 presents descriptive statistics for demographic and alcohol-related variables by sample. To participate in the present studies, participants had to be registered on MTurk and voluntarily respond to the study posting on the MTurk website. Inclusion criteria were as follows: (1) currently live in the United States, (2) ≥25 years of age (i.e., older than the "typical" college student), and (3) consumed alcohol at least once in the past month (the same criterion has been used by previous studies on PBS use among college students [e.g., Martens et al., 2007]). Those who were eligible completed a brief online survey for $0.40, which is commensurate with similar tasks on MTurk. Notably, few (~14%) MTurk participants report that MTurk is their primary source of income (Paolacci, Chandler, & Ipeirotis, 2010). Participants who failed one of two attention checks were excluded from the analyses. Also, participants who were ≥65 years of age were excluded from the analyses because geriatric adults drink less than younger adults (Substance Abuse and Mental Health Services Administration [SAMHSA], 2014) and have different health recommendations for their drinking (National Institute on Alcohol Abuse and Alcoholism [NIAAA], 2005). Ethical approval for both studies was obtained from the IRB at the University of Texas at El Paso.

2.2. Measures

2.2.1. PBS

The PBSS-20 (Treloar et al., 2015) is a 20-item measure that assesses frequency of PBS use and was previously described in further detail. Participants are instructed to rate the degree to which they engage in PBS when using alcohol or "partying" on a 6-point response scale ranging from 1 (*Never*) to 6 (*Always*). Subscales are created by summing their respective items.

2.2.2. Alcohol use

A Quantity-Frequency Index (Cahalan, Cisin, & Crossley, 1969) was used to assess alcohol use in the past three months. Frequency of alcohol use, number of alcoholic drinks consumed on a typical drinking occasion, and number of alcoholic drinks consumed on the heaviest drinking occasion were assessed with one question each. Thus, three separate alcohol use outcomes are derived from this measure. Participants were shown a chart that defined a standard drink based on different container sizes and types of alcohol consumed. Response options for the frequency item were on a response scale that ranged from 1 (*Less than once a month*) to 6 (*Once a day or more*) and response options for the two quantity items ranged from 1 (*1 drink*) to 31 (*More than 30 drinks*).

2.2.3. Alcohol-related problems

Alcohol-related problems during the past year were assessed using the Short Inventory of Problems +6 (SIP +6; Soderstrom et al., 2007), which consists of the 15-item Short Inventory of Problems (SIP; Miller, Tonigan, & Longabaugh, 1995) plus 6 additional items that are related to injury. Each item (e.g., "I have taken foolish risks because of drinking") is responded to on a 5-point response scale ranging from 0 (*Never*) to 4 (*Daily or almost daily*). The 21 items were summed to create a total score. Internal consistency estimates of the SIP +6 across all three samples were high ($\alpha = 0.91–0.93$).

3. Results

3.1. Sample 1 (n = 360): exploratory factor analysis

An exploratory factor analysis (EFA) was conducted using M*plus* Version 8 (Muthén & Muthén, 1998–2017) with maximum likelihood estimation (ML) and oblique geomin rotation. To determine the number of factors to retain, an accompanying parallel analysis (Horn, 1965) was conducted and the eigenvalues derived from the sample data were compared to the mean eigenvalues derived from 50 random data sets. The number of factors retained was the number of eigenvalues from the sample data that were larger than the mean eigenvalues from the random data. Furthermore, the PBSS-20 items were developed for use with college students, and thus item reduction procedures were used as some items may not perform well for use with a more general population of adult drinkers. Based on the recommendations of Tabachnick and Fidell (2007), items with low loadings (< 0.32) or relatively strong loadings (≥0.32) onto two or more factors (i.e., crossloadings) were dropped from further analyses. Initial EFAs were conducted to determine which (if any) items should be dropped and a final EFA was conducted with the set of remaining items.

The initial EFAs resulted in dropping 5 low-loading or crossloading items based on the criteria recommended by Tabachnick and Fidell. However, another item nearly met the crossloading criterion (i.e., 0.32 loading onto one factor and 0.31 loading onto another factor) and we made the decision to also drop this item. A final EFA was subsequently conducted excluding these 6 items. The items that were dropped were Items 3, 4, 10, 11, 15 and 16. The fourth mean eigenvalue from the random data was larger than the fourth eigenvalue from the sample data of the final EFA and thus three factors were retained with the qualification that these three factors were interpretable. The eigenvalues for the first three factors were 4.670, 1.431, and 1.255, respectively, accounting for 52.54% of the variance in the items. Factors I, II, and III (see Table 2) are largely consistent with the Serious Harm Reduction, Manner of Drinking, and Stopping/Limiting Drinking factors, respectively, found among college students (e.g., Martens et al., 2005, 2007; Treloar et al., 2015). However, two Serious Harm Reduction items (Items 17 and 20) loaded the strongest onto the Manner of Drinking factor in the present study. The unstandardized item loadings onto each of the three factors and the descriptive statistics for each item are presented in Table 2.

Table 2
Unstandardized factor loadings and item descriptives of the PBSS-20 among Sample 1.

Item	Factor loadings			*M* (*SD*)
	I	II	III	
1. Use a designated driver	**0.63***	−0.13	0.01	4.66 (1.75)
2. Determine not to exceed a set number of drinks	0.02	0.27*	**0.41***	4.01 (1.64)
5. Avoid drinking games	−0.06	**0.60***	0.16*	4.37 (1.77)
6. Leave the bar/party at a predetermined time	0.10	−0.01	**0.81***	3.70 (1.62)
7. Make sure that you go home with a friend	**0.61***	0.00	0.16*	4.19 (1.74)
8. Know where your drink has been at all times	**0.37***	0.23*	−0.02	5.13 (1.29)
9. Stop drinking at a predetermined time	−0.01	0.12	**0.73***	3.74 (1.53)
12. Avoid mixing different types of alcohol	−0.04	**0.55***	0.13*	4.04 (1.68)
13. Drink slowly, rather than gulp or chug	0.10	**0.64***	0.01	4.54 (1.36)
14. Avoid trying to keep up or out-drink others	0.02	**0.59***	−0.00	4.64 (1.56)
17. Avoid combining alcohol with marijuana	0.14	**0.48***	−0.09	5.02 (1.70)
18. Avoid "pregaming" (i.e., drinking before going out)	0.04	**0.70***	0.02	4.23 (1.77)
19. Make sure you drink with people who can take care of you if you drink too much	**0.44***	0.19	0.11	4.98 (1.26)
20. Eat before or during drinking	0.16	**0.42***	−0.07	5.11 (0.89)

Note. Items are numbered in the order that they appear in the PBSS-20. Factor loadings are unstandardized. Factor loadings in boldface represent the factor which the item loaded onto the strongest. PBSS-20 = Protective Behavioral Strategies Scale-20; I = Serious Harm Reduction; II = Manner of Drinking; III = Stopping/Limiting Drinking.

* $p < .05$.

3.2. Sample 2 (n = 339): confirmatory factor analysis, internal consistency, and concurrent validity

3.2.1. Confirmatory factor analysis

Confirmatory factor analysis (CFA) was used to examine model fit of the three-factor structure of the PBSS-20 determined by the EFA. The CFA was conducted using M*plus* Version 8 (Muthén & Muthén, 1998–2017) with maximum likelihood estimation with robust standard errors (MLM). All of the factor loadings were freely estimated and the factor variances were set to 1. The covariance between Items 6 and 9 was estimated based on previous CFAs of the PBSS that also estimated this parameter because of high covariance (Cadigan et al., 2015; Martens et al., 2007). Also, a relatively large modification index suggested estimating the residual covariance between Items 1 and 7 ("Use a designated driver" and "Make sure you go home with a friend", respectively). While methodologists have warned against freeing model parameters based on modification indices (MacCallum, Roznowski, & Necowitz, 1992), the reading of the two item stems provided conceptual justification for estimating this additional parameter. It seems likely that people's friends are often their designated drivers and getting a ride home from a friend is the same as going home with a friend.

According to the joint cutoff criteria suggested by Hu and Bentler (1999) (CFI ≥ 0.95, RMSEA ≤ 0.06, SRMR ≤ 0.08), the revised three-factor model of the PBSS-20 provided an adequate fit to the data (see Table 3). Both the CFI and RMSEA values did not meet the criteria suggested by Hu and Bentler, but others have suggested that CFI values close to 0.90 are indicative of a good fitting model (Blackburn, Donnelly, Logan, & Renwick, 2004) and RMSEA values of 0.08 or lower are indicative of a mediocre fitting model (MacCallum, Browne, & Sugawara, 1996). The standardized factor loadings for the revised three-factor model of the PBSS-20 and the item descriptives are presented in Table 4. Additionally, the interfactor correlations ranged from 0.55 to 0.75.

Table 3
Summary of the global fit indices for the three PBSS-20 models tested among Sample 2.

Model	SB χ^2	*df*	CFI	RMSEA [90% CI]	SRMR	AIC
Revised three-factor PBSS-20	180.19	72	0.905	0.067 [0.054, 0.079]	0.055	16,307.82
Original three-factor PBSS-20 (14 items)	204.30	72	0.884	0.074 [0.062, 0.086]	0.062	16,334.96
Original three-factor PBSS-20 (20 items)	433.99	165	0.842	0.069 [0.061, 0.077]	0.066	23,206.43

Note. PBSS-20 = Protective Behavioral Strategies Scale-20; SB χ^2 = Satorra-Bentler scaled χ^2; CFI = Comparative Fit Index; RSMSEA = root mean square error of approximation; CI = confidence interval; SRMR = standardized root mean square residual; AIC = Akaike information criterion.

Best practices in CFA procedures include testing alternative models to avoid confirmation bias (e.g., MacCallum & Austin, 2000). Perhaps the most likely alternative model is the original three-factor model of the PBSS-20 found among college students (Treloar et al., 2015). Thus, we compared our revised three-factor model of the PBSS-20 to two other models. These comparisons were made using the Akaike information criterion (AIC)—the model with the lowest AIC is the preferred model (Brown, 2014). The first comparison model was the original three-factor model of the PBSS-20 found among college students that excluded the 6 items that were dropped based on the EFAs conducted with the Sample 1 data. The only difference between the original and revised three-factor models excluding those 6 items is the two Serious Harm Reduction items that were loaded onto the Manner of Drinking factor in the revised model. The second comparison model that was tested was the original three-factor model of the PBSS-20 found among college students with all 20 of the original items included. The covariance between Items 6 and 9 was estimated in both comparison models and, similar to the revised model, modification indices for both comparison models indicated that the unique variances for Items 1 and 7 should be allowed to covary. Table 3 displays the model fit indices for all three models that were tested. As shown, the revised three-factor model had the lowest AIC and best global fit indices and was thus selected as the preferred model.

3.2.2. Internal consistency

Internal consistency of the three revised PBSS-20 subscales was assessed using Cronbach's alpha (α), with an asymptotically distribution-free (ADF) 95% CI (Maydeu-Olivares, Coffman, & Hartmann, 2007) constructed around α, and an ordinal version of α that has been found to be a less biased indicator of internal consistency for Likert-type data than α (Zumbo, Gadermann, & Zeisser, 2007). The internal

Table 4
Standardized factor loadings (standard errors) and item descriptives of the PBSS-20 among Sample 2.

Item	λ (SE)	*M* (*SD*)
Serious Harm Reduction		
1. Use a designated driver	0.31 (0.07)	4.65 (1.71)
7. Make sure that you go home with a friend	0.56 (0.05)	4.08 (1.70)
8. Know where your drink has been at all times	0.66 (0.05)	5.22 (1.28)
19. Make sure you drink with people who can take care of you if you drink too much	0.67 (0.05)	4.99 (1.29)
Stopping/Limiting Drinking		
2. Determine not to exceed a set number of drinks	0.67 (0.05)	4.01 (1.66)
6. Leave the bar/party at a predetermined time	0.66 (0.05)	3.60 (1.62)
9. Stop drinking at a predetermined time	0.81 (0.04)	3.69 (1.57)
Manner of Drinking		
5. Avoid drinking games	0.60 (0.04)	4.29 (1.81)
12. Avoid mixing different types of alcohol	0.55 (0.05)	4.05 (1.68)
13. Drink slowly, rather than gulp or chug	0.70 (0.04)	4.53 (1.30)
14. Avoid trying to keep up or out-drink others	0.51 (0.06)	4.64 (1.56)
17. Avoid combining alcohol with marijuana	0.47 (0.06)	4.85 (1.68)
18. Avoid "pregaming" (i.e., drinking before going out)	0.67 (0.04)	4.19 (1.69)
20. Eat before or during drinking	0.45 (0.06)	5.09 (0.98)

Note. Items are numbered in the order that they appear in the PBSS-20. PBSS-20 = Protective Behavioral Strategies Scale-20.

consistency estimates for the Stopping/Limiting Drinking (α = 0.79, ADF 95% CI [0.74, 0.84]; ordinal α = 0.83) and Manner of Drinking (α = 0.76, ADF 95% CI [0.71, 0.80]; ordinal α = 0.81) subscales were acceptable, but α for Serious Harm Reduction subscale was just below the commonly accepted threshold of 0.70 (α = 0.66, ADF 95% CI [0.60, 0.73]; ordinal α = 0.74).

3.2.3. Concurrent validity

Concurrent validity was assessed by computing both Pearson's product-moment correlation and Spearman's rank-order correlation between the three PBSS-20 subscales and the alcohol-related variables (see Table 5). Spearman's correlation—a nonparametric test of association—was computed because the alcohol use variables are count data that are not normally distributed. All three subscales demonstrated statistical, negative correlations with all of the alcohol-related variables except for frequency of alcohol use—the two types of correlations yielded similar results.

Table 5
Bivariate correlations between the PBSS-20 subscales and the alcohol-related variables among Sample 2.

Variable	Frequency	Typical quantity	Heaviest quantity	SIP +6
Pearson's product-moment correlation				
Manner of Drinking	−0.06	−0.33**	−0.43**	−0.34**
Stopping/Limiting Drinking	−0.09	−0.23**	−0.38**	−0.26**
Serious Harm Reduction	−0.07	−0.16**	−0.18**	−0.26**
Spearman's rank-order correlation				
Manner of Drinking	−0.07	−0.34**	−0.41**	−0.40**
Stopping/Limiting Drinking	−0.09	−0.26**	−0.44**	−0.30**
Serious Harm Reduction	−0.06	−0.11*	−0.17**	−0.23**

Note. Frequency = frequency of alcohol use in the past three months. Typical quantity = number of alcoholic drinks consumed on a typical drinking occasion in the past three months; Heaviest quantity = number of alcoholic drinks consumed on the heaviest drinking occasion in the past three months; SIP +6 = Short Inventory of Problems +6.
* $p < .05$.
** $p < .01$.

3.3. Sample 3 (n = 338): gender invariance

Tests of measurement invariance of the revised three-factor model of the PBSS-20 across gender were conducted using CFA procedures in M*plus* Version 8 (Muthén & Muthén, 1998–2017) with MLM estimation. All of the factor loadings were freely estimated and the factor variances were set to 1. The covariance between Items 1 and 7 and Items 6 and 9 were estimated. Prior to the measurement invariance analyses, the revised three-factor model was tested separately for females and males using Hu and Bentler's (1999) joint cutoff criteria. Next, measurement invariance across gender was tested using the following hierarchy: configural, metric, and scalar invariance. Configural invariance is said to occur when the pattern of factor loadings is the same across groups (Meredith, 1993). Metric invariance is said to occur when the factor loadings do not statistically differ across groups. Scalar invariance is said to occur when the latent item intercepts do not differ across groups.

The configural invariance model was tested using Hu and Bentler's criteria. The metric and scalar invariance models were compared to the appropriate less restrictive model using the difference test for the Satorra-Bentler scaled χ^2 statistic (Satorra & Bentler, 2010) and ΔCFI such that a χ^2 difference test that is not statistically significant and ΔCFI ≥ |0.01| (Cheung & Rensvold, 2002) are indicative of invariance. If metric and/or scalar invariance could not be established, partial metric and/or scalar invariance models were tested in which modification indices were used to identify noninvariant item parameters that were then freely estimated across gender (Byrne, Shavelson, & Muthén, 1989). To our knowledge, there are no suggested limits on the number of noninvariant item parameters to free in partial invariance models. Therefore, we established an arbitrary a priori limit of one-third of the items (i.e., 4 out of the 14 items), as allowing more than one-third of the items to be freely estimated across gender seemed excessive. Equality constraints were relaxed one by one on the basis of modification indices until the partial invariance model met the χ^2 difference test and ΔCFI criteria or the one-third limit was reached.

The revised three-factor model of the PBSS-20 provided an adequate fit to the data for both females and males (see Table 6). Given the above, we proceeded to the measurement invariance analyses. Table 6 provides the model fit indices and comparisons (if applicable) for each of the measurement invariance models tested. As shown, the configural invariance model adequately described the data. We then moved on to test metric invariance with the configural invariance model as the comparison. A statistically significant χ^2 difference test and ΔCFI >

Table 6
Measurement invariance and global fit indices of the PBSS-20 across gender among Sample 3.

Model	SB χ^2	ΔSB χ^2	*df*	Δ*df*	Δ*p*	CFI	ΔCFI	RMSEA [90% CI]	ΔRMSEA	SRMR
Female	116.83		72			0.942		0.054 [0.036, 0.072]		0.058
Male	103.21		72			0.904		0.058 [0.030, 0.082]		0.073
Configural	220.07		144			0.931		0.056 [0.041, 0.070]		0.064
Metric	262.30	46.38	158	14	sig	0.906	−0.025	0.062 [0.049, 0.076]	0.006	0.085
Partial metric	236.81	16.72	155	22	non-sig	0.926	−0.005	0.056 [0.041, 0.070]	0.000	0.074
Partial scalar	262.26	28.00	164	20	sig	0.911	−0.015	0.060 [0.046, 0.073]	0.004	0.087

Note. Values of ΔSB χ^2 represent the Satorra-Bentler scaled difference test statistic where Δ*df* and Δ*p* are the degrees of freedom and *p*-value associated with that test, respectively. PBSS-20 = Protective Behavioral Strategies Scale-20; SB χ^2 = Satorra-Bentler scaled χ^2; CFI = Comparative Fit Index; RSMSEA = root mean square error of approximation; CI = confidence interval; SRMR = standardized root mean square residual; sig = significant; non-sig = nonsignificant.

|0.01| suggested lack of metric invariance. However, a partial metric invariance model in which the loadings for Items 7, 8, and 9 were freely estimated across gender met the invariance criteria. Having established partial metric invariance, we continued by testing a partial scalar invariance model in which the intercepts for Items 7, 8, and 9 were also freely estimated across gender given that these items demonstrated noninvariance in the metric invariance models (Millsap, 2011). This partial scalar invariance model did not meet the invariance criteria (not reported in Table 6) nor was the invariance criteria met after relaxing another equality constraint on the intercept of Item 19 (reported in Table 6 as Partial Scalar Model). Thus, partial scalar invariance could not be established as 4 item intercepts were freely estimated across gender and the invariance criteria was not met.

4. Discussion

The present studies are the first to our knowledge to conduct a psychometric evaluation of the PBSS-20 among a more general population of U.S. adult drinkers. Consistent with most psychometric analyses of the PBSS and PBSS-20 among college students (e.g., Martens et al., 2005; Martens et al., 2007; Treloar et al., 2015), the present studies found support for a three-factor model of the PBSS-20 representing Serious Harm Reduction, Manner of Drinking, and Stopping/Limiting Drinking strategies. However, item reduction procedures suggested excluding 2 Serious Harm Reduction and 4 Stopping/Limiting Drinking items that did not perform well among Sample 1. Because the items of the PBSS and PBSS-20 were created for use with college students, we eliminated these items as they may not adequately reflect PBS use among a more general population of U.S. adult drinkers, which resulted in 14 remaining items. Furthermore, the results of the EFA suggested that two Serious Harm Reduction items load onto the Manner of Drinking factor. These two items, “Avoid combining alcohol with marijuana” and “Eat before or during drinking”, were added to the Serious Harm Reduction subscale in the PBSS-20 to improve its content validity (Treloar et al., 2015). Conceptually, however, both of these items seem to be more consistent with Manner of Drinking strategies. While using alcohol with marijuana or on an empty stomach may lead to higher levels of intoxication, these behaviors do not imply limiting the experience of serious harm in the same way that the other Serious Harm Reduction items do (e.g., “Use a designated driver”). In contrast, limiting intoxication is the theme among the Manner of Drinking items (e.g., “Avoid drinking games”) and Items 17 and 20 are consistent with this theme. The revisions to the PBSS-20 described above were further supported among Sample 2 as the revised three-factor model provided a better fit to the data than the original three-factor model with both 14 and 20 items.

We also examined the internal consistency and the concurrent validity of the three revised PBSS-20 subscales with the Sample 2 data and measurement invariance across gender with the Sample 3 data. Two of the three subscales demonstrated acceptable internal consistency estimates, but there was some evidence for low internal consistency of the Serious Harm Reduction subscale. While the expanded Serious Harm Reduction subscale of the PBSS-20 has demonstrated adequate internal consistency estimates (α = 0.82–0.86; Treloar et al., 2015), the Serious Harm Reduction subscale of the PBSS has demonstrated low internal consistency estimates (e.g., α = 0.59; Martens et al., 2007). Given that α is a function of the number of items in a scale (e.g., Morera & Stokes, 2016), it is not surprising that the internal consistency estimate for the revised Serious Harm Reduction subscale (4 items) among Sample 2 was in between the 3- and 8-item versions of this subscale. Also, more frequent use of all three types of strategies was associated with consuming fewer alcoholic drinks on both a typical and the heaviest drinking occasion in the past three months and with experiencing fewer alcohol-related problems over the past year.

These findings are consistent with the PBS literature as a whole and the subset of studies that have used the PBSS and have found negative correlations between the three subscales and alcohol-related variables (Pearson, 2013). Notably, PBS use was not associated with the frequency of alcohol use among Sample 2, which was expected given that the PBSS-20 uses the narrower definition of PBS. Behaviors that are used *while* drinking cannot influence the frequency of drinking. Lastly, we found a lack of measurement invariance of the revised PBSS-20 across gender as partial scalar invariance could not be established. This suggests that the items and factors of the PBSS-20 are conceptually different across gender. Studies with college student samples have also found a lack of measurement invariance across gender for both the PBSS and the PBSS-20 (Treloar et al., 2014, 2015, respectively). Treloar et al. (2014) provide recommendations for addressing the lack of measurement invariance across gender.

4.1. Limitations

Several limitations to the present studies should be acknowledged. First, although MTurk samples are more geographically diverse than standard internet samples, MTurk participants differ from the general U.S. population (e.g., MTurk participants are younger, more educated, and include a greater number of females; Ross, Irani, Silberman, Zaldivar, & Tomlinson, 2010). Thus, the external validity of the present studies may be limited, such that the findings of the present study may not generalize to other samples of U.S. adult drinkers. MTurk samples, however, are more diverse than typical college student samples (Buhrmester et al., 2011), and the present studies represent an important first step in extending research on PBS beyond strictly college student samples. Similarly, convenience sampling was used given that participants voluntarily responded to the study posting on the MTurk website, which may have resulted in a biased sample. Second, the present studies did not assess whether participants were currently enrolled in college. It may be that some people in the sample were enrolled in college and potentially exposed to a college drinking environment. Third, given the cross-sectional nature of the present study, only correlational relationships between PBS and alcohol use and related problems could be established. Lastly, although 6 items from the PBSS-20 were eliminated among Sample 1, Samples 2 and 3 completed the original 20-item version of the PBSS-20. Thus, we are uncertain as

to how responding to the 6 eliminated items may have influenced responses to the set of 14 items that was retained in the revised version of the PBSS-20.

4.2. Conclusions

Despite these limitations, the present studies are the first to examine the psychometric properties of the PBSS-20 among U.S. adult drinkers that were not recruited from a college or university and thus are an important step forward in extending research on PBS. While the three-factor model of the PBSS-20 found among college students was largely supported in the present studies, the results suggest revising the measure for use with adult drinkers. These revisions include deleting 6 items and using 2 Serious Harm Reduction items (Items 17 and 20) as Manner of Drinking items. Further psychometric analyses of the PBSS-20 with non-college student samples are needed to support revisions to the factor structure suggested by the present studies and to increase the generalizability of the findings given that MTurk samples may not be representative of the general U.S. population. The present studies are also the first to examine the association between PBS use and alcohol-related outcomes after having first verified the three subscales among the study samples as well as the first to examine measurement invariance across gender among adult drinkers.

Additional research on the validity of the PBSS in adult populations is important as refinement or development of measures of PBS use among non-college student populations is much needed. Future research might focus on developing items via focus groups with non-college student samples. Given that both the PBSS and PBSS-20 were developed for use with college students, the PBSS and PBSS-20 may not include items that represent important PBS for the more general population of U.S. adult drinkers. Testing Items 17 and 20 as Manner of Drinking items among college students may also be an interesting direction for future research, especially because the present studies are the first to attempt to replicate the psychometric properties of the PBSS-20 among any sample. As noted previously, Items 17 and 20 seem to fit better with the Manner of Drinking factor and there has been some lack of fit of the three-factor model of the PBSS-20 among college students (Treloar et al., 2015). Future research can also consider assessing measurement invariance across age and drinking patterns for the PBSS-20 among non-college student samples. While the samples in the present studies included people of diverse ages and drinking patterns, these analyses were not conducted with data from the present studies because conducting further measurement invariance analyses may have risked capitalizing on chance characteristics of the data. Finally, research establishing the relationship between PBS use and alcohol use and related problems over time would be essential for establishing predictive validity. Regardless of the direction of future research, the extension of PBS research to other populations is important because PBS may useful in reducing the public health burden of alcohol misuse among the general population of adult drinkers living in the U.S.

Declarations of interest

None.

References

Barnett, N. P., Murphy, J. G., Colby, S. M., & Monti, P. M. (2007). Efficacy of counselor vs. computer-delivered intervention with mandated college students. *Addictive Behaviors, 32*, 2529–2548.

Blackburn, R., Donnelly, J. P., Logan, C., & Renwick, S. J. D. (2004). Convergent and discriminative validity of interview and questionnaire measures of personality disorders in mentally disordered offenders: A multitrait-multimethod analysis using confirmatory factor analysis. *Journal of Personality Disorders, 18*, 129–150.

Brown, T. A. (2014). *Confirmatory factor analysis for applied research.* New York, NY: Guilford.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science, 6*, 3–5.

Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin, 105*, 456–466.

Cadigan, J. M., Weaver, C. C., McAfee, N. W., Herring, T. E., & Martens, M. P. (2015). A confirmatory factor analysis of the Protective Behavioral Strategies Scale among OEF/OIF veterans. *Addictive Behaviors, 51*, 127–130.

Cahalan, D., Cisin, I. H., & Crossley, H. M. (1969). *American drinking practices.* New Brunswick, NJ: Rutgers Center of Alcohol Studies.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*, 233–255.

Hawkins, E. J., Lapham, G. T., Kivlahan, D. R., & Bradley, K. A. (2010). Recognition and management of alcohol misuse in OEF/OIF and other veterans in the VA: A cross-sectional study. *Drug and Alcohol Dependence, 109*, 147–153.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30*, 179–185.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*, 1–55.

Larimer, M. E., Lee, C. M., Kilmer, J. R., Fabiano, P. M., Stark, C. B., Geisner, I. M., ... Neighbors, C. (2007). Personalized mailed feedback for college drinking prevention: A randomized clinical trial. *Journal of Consulting and Clinical Psychology, 75*, 285–293.

MacCallum, R. C., & Austin, J. T. (2000). Applications of structural equation modeling in psychological research. *Annual Review of Psychology, 51*, 201–226.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*, 130–149.

MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model modifications in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin, 111*, 490–504.

Martens, M. P., Ferrier, A. G., Sheehy, M. J., Corbett, K., Anderson, D. A., & Simmons, A. (2005). Development of the Protective Behavioral Strategies Survey. *Journal of Studies on Alcohol, 66*, 698–705.

Martens, M. P., Pederson, E. R., LaBrie, J. W., Ferrier, A. G., & Cimini, M. D. (2007). Measuring alcohol-related protective behavioral strategies among college students: Further examination of the Protective Behavioral Strategies Scale. *Psychology of Addictive Behaviors, 21*, 307–315.

Maydeu-Olivares, A., Coffman, D. L., & Hartmann, W. M. (2007). Asymptotically distribution-free (ADF) interval estimation of coefficient alpha. *Psychological Methods, 12*, 157–176.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58*, 525–543.

Miller, W. R., Tonigan, J. S., & Longabaugh, R. (1995). In M. E. Mattson (Vol. Ed.), *Project MATCH monograph series: . Vol. 4*.

Millsap, R. E. (2011). *Statistical approaches to measurement invariance.* New York, NY: Routledge.

Morera, O. F., & Stokes, S. M. (2016). Coefficient α as a measure of test score reliability: Review of 3 popular misconceptions. *American Journal of Public Health, 106*, 458–461.

Murphy, J. G., Dennhardt, A. A., Skidmore, J. R., Borsari, B., Barnett, N. P., Colby, S. M., & Martens, M. P. (2012). A randomized controlled trial of a behavioral economic supplement to brief motivational interventions for college drinking. *Journal of Consulting and Clinical Psychology, 80*, 876–886.

Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide* (8th ed.). Los Angeles, CA: Muthén & Muthén.

National Institute on Alcohol Abuse and Alcoholism (NIAAA) (2005). *Helping patients who drink too much: A clinician's guide.* Rockville, MD: NIAAA.

Novik, M. G., & Boekeloo, B. O. (2011). Dimensionality and psychometric analysis of an alcohol protective behavioral strategies scale. *Journal of Drug Education, 41*, 65–78.

Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision making, 5*, 411–419.

Pearson, M. R. (2013). Use of alcohol protective behavioral strategies among college students: A critical review. *Clinical Psychology Review, 33*, 1025–1040.

Prince, M. A., Carey, K. B., & Maisto, S. A. (2013). Protective behavioral strategies for reducing alcohol involvement: A review of the methodological issues. *Addictive Behaviors, 38*, 2343–2351.

Ross, J., Irani, L., Silberman, M. S., Zaldivar, A., & Tomlinson, B. (2010). Who are the crowdworkers?: Shifting demographics in Mechanical Turk. *CHI EA '10: Proceedings of the 28th of the International Conference Extended Abstracts on Human Factors in Computing Systems* (pp. 2863–2872). New York, NY, USA: ACM.

Satorra, A., & Bentler, P. M. (2010). Ensuring positiveness of the scaled difference chi-square test statistic. *Psychometrika, 75*, 243–248.

Soderstrom, C. A., DiClemente, C. C., Dischinger, P. C., Hebel, J. R., McDuff, D. R., Auman, K. M., & Kufera, J. A. (2007). A controlled trial of brief intervention versus brief advice for at-risk drinking trauma center patients. *Journal of Trauma and Acute Care Surgery, 62*, 1102–1112.

Substance Abuse and Mental Health Services Administration (SAMHSA) (2014). *Results from the 2013 National Survey on Drug Use and Health: Summary of national findings.* Rockville, MD: SAMHSA.

Sugarman, D. E., & Carey, K. B. (2007). The relationship between drinking control strategies and college student alcohol use. *Psychology of Addictive Behaviors, 21*, 338–345.
Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics.* Boston, MA: Allyn and Bacon.
Treloar, H., Martens, M. P., & McCarthy, D. M. (2015). The Protective Behavioral Strategies Scale-20: Improved content validity of the Serious Harm Reduction subscale. *Psychological Assessment, 27*, 340–346.
Treloar, H. R., Martens, M. P., & McCarthy, D. M. (2014). Testing measurement invariance of the Protective Behavioral Strategies Scale in college men and women. *Psychological Assessment, 26*, 307–313.
Zumbo, B. D., Gadermann, A. M., & Zeisser, C. (2007). Ordinal versions of coefficients alpha and theta for Likert rating scales. *Journal of Modern Applied Statistical Methods, 6*, 21–29.