Received 12/19/17
Revised 03/01/18
Accepted 03/27/18
DOI: 10.1002/jcad.12231

Systematic Evaluation of Psychometric Characteristics of the Michigan Alcoholism Screening Test 13-Item Short (SMAST) and 10-Item Brief (BMAST) Versions

Amelia Minnich, Bradley T. Erford, Gerta Bardhoshi, Zümra Atalay, Catharine Y. Chang, and Lauren A. Muller

The psychometric properties of the Michigan Alcohol Screening Test (MAST) 13-item short (SMAST) and 10-item brief (BMAST) versions were aggregated and synthesized across 40 and 21 studies, respectively. Results for reliability, validity, and nonclinical descriptive statistics were reported separately for the SMAST and BMAST, and implications for counseling practice and research were discussed.

Keywords: Michigan Alcohol Screening Test, MAST, SMAST, BMAST, alcohol use disorder

The Michigan Alcoholism Screening Test (MAST; Selzer, 1971) is a frequently used self-report screening instrument created to detect symptoms of problematic drinking and alcohol use disorder (Minnich, Erford, Bardhoshi, & Atalay, 2018; Shields, Howell, Potter, & Weiss, 2007). Brief versions of the MAST are available, including the 13-item Short MAST (SMAST; Selzer, Vinokur, & Van Rooijen, 1975), 10-item Brief MAST (BMAST; Pokorny, Miller, & Kaplan, 1972), and nine-item and 10-item Malmo modification of the MAST (Mm-MAST; Isacsson, Hanson, Janzon, Lindell, & Steen, 1987) versions. In addition, full-length geriatric (MAST-G; Johnson-Greene, McCaul, & Roger, 2009) and short geriatric (SMAST-G; Blow, Gillespie, Barry, Mudd, & Hill, 1998) versions and a veterans' version (VAST; Magruder-Habib, Stevens, & Ailing, 1993) are available. The original MAST, its variations, and short forms have been extensively researched and used in more than 25 countries and in a variety of populations, languages, and settings (Shields et al., 2007).

Minnich et al. (2018) provided a psychometric synthesis of 103 MAST publications and derived an aggregated internal consistency (Kuder–Richardson Formula 20 [KR-20]) estimate of .85, robust internal and external validity, and a diagnostic accuracy rate close to 80%. They also noted mean score differences between men and women that may have implications for differential diagnostic accuracy. All versions of the MAST use criterion-referenced interpretation and cutoff scores. Now that extensive study of the MAST psychometric properties has been conducted, we turn our attention to the brief MAST versions to determine their efficacy for use in counseling practice and research. Thus, the purpose of this article was to identify published studies that explored the psychometric properties of the short (SMAST) and brief (BMAST) versions of the MAST and to aggregate the results to better understand the overall psychometric characteristics of these two popular shortened versions and their utility in counseling practice and research.

The SMAST is an abbreviated version of the original MAST (Selzer et al., 1975). The scoring method for the 13-item SMAST (Items 1, 3, 5, 6, 8, 9, 11, 14, 16, 20, 21, 24, and 25 of the original MAST; see Table 1) is 1 point for each response indicating alcohol use, in contrast to the weighted scoring method often used with the MAST or BMAST. Scores range from 0 to 13, and the suggested cutoff score is 3 for persons with alcohol use disorder, whereas a score of 2 is suggestive of persons with alcohol use disorders. To examine its psychometric properties, Selzer et al. (1975) administered the SMAST to 501 male drivers: 273 drivers renewing their

Amelia Minnich, Education Specialties Department, Loyola University Maryland; **Bradley T. Erford** and **Lauren A. Muller,** Department of Human and Organizational Development, Vanderbilt University; **Gerta Bardhoshi,** Rehabilitation and Counselor Education Department, University of Iowa; **Zümra Atalay,** Counseling Department, MEF University, Istanbul, Turkey; **Catharine Y. Chang,** Counseling and Psychological Services Department, Georgia State University. Amelia Minnich is now at St. Paul's Lutheran School, Glen Burnie, Maryland. Correspondence concerning this article should be addressed to Bradley T. Erford, Department of Human and Organizational Development, Peabody College, Vanderbilt University, 230 Appleton Place, Nashville, TN 37203 (email: bradley.t.erford@vanderbilt.edu).

TABLE 1

Item-Scoring Correspondence for the Original Michigan Alcoholism Screening Test (MAST), Short Michigan Alcoholism Screening Test (SMAST), and Brief Michigan Alcoholism Screening Test (BMAST)

Item Number				Responses and Weights			
MAST	SMAST	BMAST	Item	MAST-U	MAST-W	SMAST	BMAST
1	1	1	Do you feel you are a normal drinker?	No	No (2)	No	No (2)
2			Have you ever awakened the morning after some drinking the night before and found that you could not remember a part of the evening before?	Yes	Yes (2)		
3	2		Does your wife/husband (or parents) ever worry or complain about your drinking?	Yes	Yes (1)	Yes	
4			Can you stop drinking without a struggle after one or two drinks?	No	No (2)		
5	3		Do you ever feel bad after your drinking?	Yes	Yes (1)	Yes	
6	4	2	Do friends or relatives think you are a normal drinker?	No	No (2)	No	No (2)
7[a]			Do you ever try to limit your drinking to certain times of the day or certain places?	Yes	Yes (0)		
8	5		Are you always able to stop drinking when you want to?	No	No (2)	No	
9[b]	6	3	Have you ever attended a meeting of Alcoholics Anonymous?	Yes	Yes (5)	Yes	Yes (5)
10			Have you gotten into fights when drinking?	Yes	Yes (1)		
11	7		Has drinking ever created problems between you and your wife/husband?	Yes	Yes (2)	Yes	
12			Has your wife/husband (or other family member) ever gone to anyone for help about your drinking?	Yes	Yes (2)		
13		4	Have you ever lost friends or girlfriends/boyfriends because of drinking?	Yes	Yes (2)		Yes (2)
14	8	5	Have you ever gotten into trouble at work because of drinking?	Yes	Yes (2)	Yes	Yes (2)
15			Have you ever lost a job because of drinking?	Yes	Yes (2)		
16	9	6	Have you ever neglected your obligations, your family, or your work for 2 or more days in a row because you were drinking?	Yes	Yes (2)	Yes	Yes (2)
17			Do you ever drink before noon?	Yes	Yes (1)		
18			Have you ever been told you have liver trouble? Cirrhosis?	Yes	Yes (2)		
19		7	Have you ever had delirium tremens, severe shaking, heard voices, or seen things that weren't there after heavy drinking?	Yes	Yes (2)		Yes (2)
20[c]	10	8	Have you ever gone to anyone for help about your drinking?	Yes	Yes (5)	Yes	Yes (5)
21[c]	11	9	Have you ever been in a hospital because of drinking?	Yes	Yes (5)	Yes	Yes (5)
22[c]			Have you ever been a patient in a psychiatric hospital or on a psychiatric ward of a general hospital where drinking was a part of the problem?	Yes	Yes (2)		
23[c]			Have you ever been seen at a psychiatric or mental health clinic or gone to a doctor, social worker, or clergyman for help with an emotional problem in which drinking has played a part?	Yes	Yes (2)		
24	12		Have you ever been arrested, even for a few hours, because of drunk behavior?	Yes	Yes (2)	Yes	
25	13	10	Have you ever been arrested for drunk driving after drinking?	Yes	Yes (2)	Yes	Yes (2)

Note. Number in parentheses for MAST-W and BMAST indicates number of points for the indicated response. MAST-U = MAST unit scoring method (1 point for the indicated response); MAST-W = MAST weighted scoring method.
[a]Item 7 was dropped to compose the 24-item version of the MAST. [b]Should be attendance for personal problems, not as an observer or supportive friend or family member. [c]For Items 20–23, do not include this current outpatient or hospital episode, if applicable.

licenses or participating in court-mandated driver safety school, and 228 drivers receiving inpatient or outpatient treatment for alcohol use disorder. The SMAST scores displayed less than adequate internal consistency (KR-20 < .80; Erford, 2013) for the license renewal and safety school group (α = .76) and for the inpatient and outpatient group (α = .78). However, when combined, the total sample score yielded a high reliability coefficient of .93. In a meta-analysis of the MAST and SMAST, there were no significant differences in score reliability (Shields et al., 2007).

Pokorny et al. (1972) introduced a 10-item BMAST, which included items of the MAST that were considered most discriminating in identifying persons with alcohol use disorder and used a weighted scoring system (see Table 1). A cutoff score of 6 was recommended. Pokorny et al. found that the 10-item BMAST (i.e., Items 1, 6, 9, 13, 14, 16, 19, 20, 21, and 25 of the original MAST) was as effective as the full 25-item MAST in detecting persons with alcohol use disorder. The BMAST uses a weighted scoring system to give greater emphasis to those items considered most predictive of alcohol dependence. However, the weighted scores for each version on the full MAST and the BMAST were highly correlated (.95 to .99; Pokorny et al., 1972), although Porkony et al.'s study had been criticized for its small, relatively restricted clinical sample of individuals with alcohol use disorder who were receiving inpatient psychiatric treatment in a Veterans Administration hospital (Selzer et al., 1975).

These brief versions of the MAST have been used for nearly 40 years, and reports of psychometric veracity now number in the dozens for both versions. Thus, the time is

right to synthesize this existing evidence so that counselors and researchers can make informed judgment on use. In the present systematic evaluation, we conducted a thorough review of the SMAST and BMAST since their original development, with several research questions addressed separately for each version:

1. What is the test–retest and internal consistency reliability of scores using the SMAST and BMAST?
   a. Does the internal consistency of the SMAST and BMAST vary across clinical and nonclinical populations?
2. How well do the SMAST and BMAST correlate with other measures of substance use?
3. In nonclinical samples, what are the reported means and standard deviations of the SMAST and BMAST?
   a. Are there gender differences in mean scores and standard deviations that have implications for diagnostic prevalence and accuracy?
4. What factor-analytic evidence exists to support the internal structure of the SMAST and BMAST?
5. What is the diagnostic accuracy of the SMAST and BMAST?

To fully review and critique the SMAST and BMAST, we will address some of the additional criticisms of the MAST in the Discussion section. Magruder-Habib, Harris, and Fraker (1982) stated that due to the lifetime occurrence focus and the dichotomous response format (i.e., yes/no), the MAST does not differentiate individuals who are currently engaging in problematic drinking from those who are in recovery. Additionally, based on this systemic review, we provide recommendations related to the unit versus weighted scoring methods.

Method

Journal articles, dissertations, and other electronically available sources such as conference proceedings that met the predetermined criteria were included in this systematic analysis. Studies included in the review (a) used the English version of BMAST (Pokorny et al., 1972) or SMAST (Selzer et al., 1975); (b) were published between 1972 and 2016; and (c) provided some type of validity, reliability, or sample mean data. Thus, we attempted to capture and analyze all available psychometric studies that used the BMAST and SMAST.

Search Strategies

The first two authors selected the articles for review by conducting a systematic review of articles indexed in PsycINFO, PsycARTICLES, ERIC, Academic Search Premier, and MEDLINE from 1972 to 2016 by using the keyword *MAST* in full text. Additional SMAST and BMAST studies were located by searching reference lists of selected studies and previous synthesis articles. Selector agreement was 98.8%, and articles that met inclusion criteria were analyzed and disaggregated by instrument version. No attempt was made to locate literature not available via electronic or hand searches.

Psychometric Variables Analyzed and Statistical Methods Used

Internal consistency (KR-20), test–retest reliability, convergent correlations with other substance use measures, structural validity (i.e., exploratory factor analysis [EFA], confirmatory factor analysis [CFA]), diagnostic validity across various cutoff scores and samples (e.g., percentage of correct classifications), sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and descriptive statistics from nonclinical samples (i.e., means and standard deviations) served as the primary variables for evaluation.

All analyzed data were independent (Erford, Savin-Murphy, & Butler, 2010); that is, samples represented in more than one article or study were used only once. Because of the dichotomous format of all the items, we used the KR-20 as the test statistic for all the internal consistency analyses. We used Pearson's *r* effect size estimates for test–retest reliability and convergent validity analyses. When combined, sample sizes were used to weight the results. Following classical test theory (Erford, 2013), we analyzed internal consistency and test–retest reliability coefficient (r_{tt}) directly after weighting. Pearson's *r*s for convergent validity comparisons were initially transformed into *z* values (½log[(1 r)/(1 – r)]; Hedges & Olkin, 1985), then weighted by sample size, summed, and averaged before the grand *z* was back-transformed to *r*. Effect sizes for Pearson's *r* were classified as .1 for small effect; .3 for medium effect; and .5 for large effect (Erford et al., 2010).

Results

Of candidate articles, 702 were identified through computerized searches and 57 more through hand searches, for a total of 759 candidate articles. Full text review of all candidate articles eliminated 582 articles that violated one or more inclusion criteria. Thus, 177 articles were accepted using the MAST or at least one of its various versions (i.e., MAST, BMAST, SMAST, VAST, Mm-MAST, MAST-G, SMAST-G). Further selection procedures were used to specify articles that only used the BMAST and SMAST versions, which resulted in a final set of 40 SMAST and 21 BMAST candidate articles accepted into the analyses. Again, selector agreement between the first two authors was 98.8% and disagreements were adjudicated by consensus.

Results Using SMAST

Reliability of scores on SMAST. A total of 13 SMAST studies with a combined sample size of 7,622 participants reported

internal consistency data. When all studies were weighted and then averaged, the mean internal consistency coefficient was .77. In the clinical samples with a combined sample size of 539, internal consistency was .67, whereas in the nonclinical samples with a large combined sample size of 5,923 participants, internal consistency was .75. Three SMAST studies of test–retest reliability were located. Two of the studies reported data for a 2-week timeframe, with a combined sample size of 355 yielding r_{tt} = .74. One study (N = 2,270) reported a 6-month test–retest reliability score of .89.

Validity of scores on SMAST. Twenty convergent (external) validity studies were reported (see Table 2), with correlation coefficients ranging from .01 to .95. Two SMAST EFA studies of structural validity were located. Both studies supported a three-factor solution. Skinner and Allen (1983) identified the three dimensions of help seeking/conflict, not normal, and family discord. No CFA studies of the SMAST were located.

Nineteen studies reported diagnostic validity for the SMAST (see Table 3). The suggested cutoff score of 3 resulted in sensitivity of .68 and specificity of .74 (j = 9, N = 2,349; where j = number of studies). Not enough studies reported PPV and NPV to provide a reliable analysis. A cutoff score of 4 actually appeared slightly more parsimonious at a sensitivity of .70 and specificity of .71, but only involved three studies and a combined sample of 1,237 participants.

Descriptive characteristics of scores on SMAST. Finally, 11 SMAST studies presented nonclinical participant descriptive data, with a combined sample size of 3,792 nonclinical participants. The total sample mean was 1.48 (SD = 1.80). Only two of the studies provided descriptive statistics for male participants (N = 302), with an average mean of 1.78 (SD = 1.69), whereas three of the studies presented descriptive data for female participants (N = 388), with an average mean of 1.36 (SD = 1.21).

TABLE 3

13-Item Short Michigan Alcoholism Screening Test Diagnostic Efficiency Summary

Cutoff	*j*	*N*	Sensitivity	Specificity
1+	3	1,050	.95	.23
2+	4	1,422	.86	.42
3+	9	2,349	.68	.74
4+	3	1,237	.70	.71
5+	3	1,309	.64	.81
6+	2	957	.62	.86
7+	2	957	.68	.84
8+	2	957	.44	.96
9+	2	957	.37	.98
10+	3	1,029	.31	.98

Note. Cutoff = cutoff score; *j* = number of studies.

Results Using BMAST

Reliability of scores on BMAST. A total of five BMAST studies (Bliss, Ogley-Oliver, Jackson, Harp, & Kaslow, 2008; Lehavot & Simoni, 2011; Underwood, 2002; Willenbring, Christensen, Spring, & Rasmussen, 1987; Zung, 1979) with a combined sample size of 1,856 participants reported internal consistency data. When all studies were weighted and then averaged, the mean internal consistency coefficient was .73. In the clinical samples (j = 2; Willenbring et al., 1987; Zung, 1979), with a combined sample size of 252, internal consistency was .82. In the nonclinical samples (j = 3; Lehavot & Simoni, 2011; Underwood, 2002; Willenbring et al., 1987), with a combined sample size of 1,526, internal consistency was .71. A single BMAST study (Bernadt, Mumford, &

TABLE 2

Convergent Validity of the 13-Item Short Michigan Alcoholism Screening Test (SMAST)

r	*j*	*N*	Comparison	Articles
.30	4	308	DAST	Buchanan, 2008; Hequembourg et al., 2008; James & Taylor, 2007; Josephson, 1999
.66	3	2,802	CAGE questionnaire	Lee & DeFrank, 1988; Nilssen et al., 1994; Saunders & Kershaw, 1980
.95	4	947	MAST	Lee & DeFrank, 1988; Selzer et al., 1975; Willenbring et al., 1987; Zung, 1979
.83	2	176	BMAST	Lee & DeFrank, 1988; Willenbring et al., 1987
.42	1	37	*ICD-10* Substance Use Disorder	Møller & Linaker, 2010
.74	1	37	Clinical Alcohol Use Scale	Møller & Linaker, 2010
.70	1	276	Father vs. child's rating of father on SMAST	Crews & Sher, 1992
.63	1	389	Mother vs. child's rating of mother on SMAST	Crews & Sher, 1992
.52	1	216	Siblings SMAST agreement of mother's drinking	Crews & Sher, 1992
.86	1	216	Siblings SMAST agreement of father's drinking	Crews & Sher, 1992
.25	1	287	AUDIT total	Barry & Fleming, 1993
.01	1	287	AUDIT Drinking Behavior	Barry & Fleming, 1993
.31	1	287	AUDIT Case Subtest	Barry & Fleming, 1993
.49	1	287	AUDIT Consequences	Barry & Fleming, 1993

Note. Twenty convergent studies were embedded in the 13 citations listed in the table. *j* = number of studies; DAST = Drug Abuse Screening Test; MAST = Michigan Alcoholism Screening Test; BMAST = Brief MAST; *ICD-10* = *International Classification of Diseases and Related Health Problems* (10th ed.); AUDIT = Alcohol Use Disorders Identification Test.

Murray, 1984; $N = 371$) of test–retest reliability was located and reported data for a 5-day timeframe, yielding $r_{tt} = .71$. One BMAST study (Pokorny et al., 1972; $N = 122$) reported unit scoring versus weighted scoring procedures, yielding a correlation of .99.

Validity of scores on BMAST. Five BMAST convergent validity studies were located, with coefficients ranging from .21 to .74. One study ($N = 1{,}381$; Lehavot & Simoni, 2011) compared the BMAST with the Brief Drug Abuse Screening Test (DAST), yielding a weak, albeit statistically significant correlation of $r = .21$. In another study ($N = 236$; Connor, Grier, Feeney, & Young, 2007), the BMAST was compared with the Alcohol Use Disorders Identification Test (AUDIT), resulting in $r = .58$. A convergent validity study of the BMAST and the SMAST (Lee & DeFrank, 1988) yielded a strong correlation of .74. Finally, two studies (combined $N = 455$; Lee & DeFrank, 1988; Tabisz et al., 1991) compared the BMAST with the CAGE questionnaire, yielding a weighted average of $r = .35$.

Two BMAST studies of structural validity were located, one of which (Connor et al., 2007) presented both EFA and CFA evidence, whereas the other (Willenbring et al., 1987) presented only EFA evidence. Both EFA studies supported a two-factor solution, accounting for about 56% of item variance. Connor et al. (2007) assessed a large sample size ($N = 3{,}179$) and presented perception of current drinking and drinking consequences dimensions. Connor et al. also tested the original one-factor model (comparative fit index [CFI] = .836) and the two-factor model (CFI = .916).

Thirteen studies reported diagnostic validity for the BMAST. The reported cutoff scores ranged from 2 to 6. When aggregated, the suggested cutoff score of 6 resulted in percentage correctly classified of .80 ($j = 4$, $N = 1{,}073$), sensitivity of .48 and specificity of .90 ($j = 12$, $N = 4{,}969$), PPV of .54 ($j = 5$, $N = 1{,}663$), and NPV of .90 ($j = 4$, $N = 1{,}358$). Overall, the optimal cutoff score appeared to be the recommended score of 6. However, at the cutoff score of 6, Corvo (2006) and Cremonte and Cherpitel (2008) reported especially low sensitivity, ranging from .00 to .04 for the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM*) Fourth Edition (American Psychiatric Association, 1994) abuse criteria.

Descriptive characteristics of scores on BMAST. Five BMAST studies (Lee & DeFrank, 1988; Lehavot & Simoni, 2011; Rospenda, Minich, Milner, & Richman, 2010; Underwood, 2002; Willenbring et al., 1987) reported descriptive data for nonclinical samples, with a combined sample size of 2,612 participants. The aggregated total mean was 1.70 ($SD = 3.31$).

Discussion

The MAST and its brief alternative versions are commonly administered alcohol use screening instruments for identifying symptoms of problematic drinking and alcohol use disorder (Shields et al., 2007). The results of this systematic analysis of the SMAST and BMAST provide moderate to robust estimates of psychometric characteristics. The following summaries discuss the results for the SMAST and BMAST versions separately.

SMAST

When evaluating results pertaining to the internal consistency of the SMAST across 13 studies ($N = 7{,}622$), we noted that they diverged from the original reliability findings reported in Selzer et al.'s (1975) SMAST study. Specifically, our results indicated a mean internal consistency coefficient of .77, a much lower coefficient than the originally reported .93 (Selzer et al., 1975). Contrary to the MAST, the average internal consistency of the SMAST was lower for clinical populations (.67, $N = 539$) compared with nonclinical populations (.75, $N = 5{,}923$). These estimates fall below the suggested criterion for both screening-level (.80) and diagnostic-level (.90) purposes (Erford, 2013). In addition, two SMAST studies of test–retest reliability provided data for a 2-week timeframe ($N = 355$, $r_{tt} = .74$). However, over a longer time period of 6 months, another study with a very large sample size ($N = 2{,}270$) reported a higher reliability score of .89, an unusual albeit welcome result. Usually, longer test–retest time intervals result in lower coefficients than shorter intervals. In view of these findings, the reliability of both the SMAST and BMAST is not adequate for screening-level and diagnostic decision-making. The full-length MAST is superior in this regard with an average KR-20 of .84 (Minnich et al., 2018).

The 20 convergent validity studies using the SMAST presented a broad range of scores, from $r = .01$ to $r = .95$ (see Table 2), with the majority of scores on convergent instruments producing moderate to large effect sizes. Not surprisingly, the highest correlations with the SMAST included the MAST ($r = .95$), the MAST Recognized Problem subscale ($r = .95$), and the Siblings SMAST Agreement of Father's Drinking subscale ($r = .86$). Only four convergent comparisons displayed correlations with small effect sizes: the AUDIT Drinking Behavior subscale ($r = .01$), the AUDIT total ($r = .25$), the Minnesota Multiphasic Personality Inventory–II (MMPI-II) Antisocial Behavior Screen ($r = .22$), and the MMPI-II Psychopathic Deviate ($r = .25$) scales.

Structural validity evidence for the SMAST was scant, as we located only two studies using EFA. These studies supported a three-factor solution, but with no additional item variance data or CFA evidence, the utility of these results is limited. Again, the actual factor structure of the SMAST is less relevant when one considers that decisions are made using the SMAST total raw score. In addition, if the overall internal consistency of the SMAST is < .80, the KR-20s for SMAST subscales are likely to be far lower, and therefore too inconsistent to yield helpful interpretations.

At the SMAST's suggested cutoff score of 3, the average sensitivity was .68 and specificity was .74 (N = 2,349). Unfortunately, a thorough analysis and interpretation of the diagnostic validity of SMAST scores was limited due to few studies reporting PPV and NPV data. A further review of the available data (N = 1,237) suggests 4 as a potential optimal cutoff score (sensitivity = .70, specificity = .71), although cutoffs of 3 and 4 each resulted in overall accuracy rates close to .71, which is about 10% lower than one can expect using the 25-item MAST total raw score (Minnich et al., 2018). Future studies exploring SMAST validity should report more complete diagnostic validity data to support reliable analyses.

Finally, 11 SMAST studies reported mean and standard deviation data for their nonclinical samples (combined N = 3,792, M = 1.48, SD = 1.80). Although few studies provided sample statistics disaggregated by gender, our analysis of the available data revealed that male participants (N = 302) had a slightly higher combined mean of 1.78 (SD = 1.69) compared with female participants (N = 388, M = 1.36, SD = 1.21). Unfortunately, both genders' combined sample sizes were small, thus limiting the generalizability of the results. Future SMAST studies should use larger sample populations. Applying a z-score analysis to these means, the recommended cutoff score of 3 on the SMAST yields a z score of 0.72 for men, which would identify about 23% of men as problem drinkers, and a z score of 1.36 for women, which would identify about 8% of women as problem drinkers. These estimates are about 50% higher than current prevalence estimates (American Psychiatric Association, 2013) and may reflect that university students often participate in nonclinical studies, and classifying university students as "normal" drinkers may skew subsequent findings. Given that the internal consistency and diagnostic validity estimates are lower for the SMAST than the MAST (Minnich et al., 2018), and that the SMAST likely identifies twice the societal base rate of men and women with alcohol problems, the full-length MAST appears to be a superior instrument to the SMAST.

BMAST

Given the fewer number of items composing the BMAST, results from five studies (N = 1,856) provided a lower internal consistency coefficient (.73) than the combined results of the MAST studies (.84; Minnich et al., 2018) and the SMAST studies (.77; see preceding section). Combined BMAST clinical samples (N = 252) had a higher internal consistency coefficient (.82) compared with nonclinical samples (N = 1,526, coefficient = .71). As with the SMAST, most of these combined estimates fall below the suggested criterion for both screening-level (.80) and diagnostic-level (.90) purposes (Erford, 2013).

Bernadt et al. (1984) was the only study (N = 122) that presented test–retest reliability for the BMAST. Using a 5-day timeframe, the study yielded a reliability coefficient of .71. Additional studies are needed that examine the temporal stability of the BMAST and to facilitate more conclusive interpretations regarding this instrument's score reliability.

Convergent validity of the BMAST with other alcohol and drug screening instruments appeared rather robust, with large effect sizes noted with both the AUDIT (r = .58) and the SMAST (r = .74). Comparisons with the CAGE questionnaire yielded a moderate effect size (j = 2, r = .35). One convergent study comparing the BMAST with the Brief DAST (which screens for misuse of drugs other than alcohol) displayed a small effect size of .21.

In terms of structural validity, two studies using EFA were located supporting a two-factor solution that accounted for 56% of item variance. This two-factor model, reported by Connor et al. (2007), consists of dimensions pertaining to perception of current drinking and drinking consequences. This solution was further supported by CFA evidence, with data fitting the two-factor model well, equivalent to the fit of the original unidimensional model. Again, interpretation of the BMAST rests on the total score, so factor structure is less important than diagnostic validity. In terms of diagnostic validity, Pokorny et al.'s (1972) recommended cutoff score of 6 was reported in 13 BMAST studies, resulting in an average sensitivity of .48, specificity of .90, PPV of .54, NPV of .90, and percentage of correct classification of .80. Thus, although specificity and NPV values were high, sensitivity and PPV values were less than adequate. Despite these results, the recommended cutoff score appeared to be optimal. An 80% correct classification rate for the BMAST is comparable to the full-length MAST results (Minnich et al., 2018).

Finally, five BMAST studies reported descriptive data (combined N = 2,612, M = 1.70, SD = 3.31), although no studies provided sample statistics disaggregated by gender. These results are similar to the SMAST's sample statistics, suggesting that both short forms of the MAST yield similar descriptive data of scores. Using the data, a cutoff score of 6 yields a z score of 1.3, so is likely to identify approximately 10% of the overall U.S. population, which is only slightly higher than the 8.5% reported overall prevalence in the *DSM Fifth Edition* (American Psychiatric Association, 2013).

Limitations and Implications for Counseling Practice

Despite the rigorous protocols used in this review, study limitations exist. Many of the comparisons were based on only one or a few studies of the SMAST or BMAST for which psychometric data were available. Although combined sample sizes were large for internal consistency, nonclinical descriptive, and diagnostic validity analyses, most of the convergent comparisons involved few studies and smaller combined samples.

The SMAST was located in twice the number of studies as the BMAST, so more confidence is warranted in those analyses. In comparison, Minnich et al. (2018) reported the full-length

MAST generated 103 accepted studies, so those results can be interpreted with a great deal of confidence and statistical power. The general conclusion when comparing the MAST results with the SMAST and BMAST results is that the MAST showed adequate internal consistency across clinical and nonclinical samples for screening purposes, whereas the SMAST and BMAST generally demonstrated internal consistencies lower than the .80 criterion expected for screening-level instruments (Erford, 2013). Additionally, whereas SMAST revealed a slightly stronger overall mean internal consistency and a higher mean internal consistency for the nonclinical population than the BMAST, the BMAST had a higher internal consistency score for the clinical population than the SMAST.

The SMAST holds an additional advantage over the BMAST, because it uses the unit scoring method rather than the weighted scoring method used by the BMAST. The unit scoring method is the simpler to score (each question is rated 0 or 1) and thus likely yields more accurate and relevant interpretations for screening and treatment. Based on Minnich et al.'s (2018) review of the MAST and the current review of the shorter versions of the MAST, it appears that the MAST is a screening inventory that yields more reliable and valid scores. Users of the SMAST and BMAST should be cautious when using these shortened inventories due to the less-than-optimal reliability scores. It is unfortunate that weighted scoring methods were derived and applied to the MAST and BMAST, because this complicates scoring and interpretation while simultaneously yielding a correlation of .99 when compared with the unit scoring method (i.e., one point per item).

The BMAST and SMAST do have the advantage of being shorter and requiring less time to administer and score when compared with the MAST, although the shortened versions do not meet the minimum criterion for internal consistency expected of both screening- and diagnostic-level instruments. In settings where time may be limited, counselors may opt for using these shortened measures, but counselors should exercise caution in reaching screening-level conclusions because of psychometric limitations. In addition, Magruder-Habib et al.'s (1982) caution about the historical retrospective of all MAST versions on "lifetime history" rather than "present circumstances" warrants mention here. Counselors must be aware that the instructions do not allow differentiation of persons with an alcohol use disorder and those in recovery.

Because limited studies are available on the SMAST and BMAST, counselors need to take appropriate precautions when administering these versions and carefully interpret results while considering the lack of multicultural and age validation. Future studies should consider including samples and reporting response differences among various racial/ethnic, gender, or age groups. This will allow for the SMAST and BMAST results to be more generalizable across populations. For example, on the SMAST and BMAST, men tended to score higher than women overall, but given the limited number of studies that reported gender-differentiated means, additional research that examines gender differences in responses could be additive to existing literature. Reporting sample means and standard deviations for adult community samples in addition to university undergraduate samples would be beneficial, because the latter may not be representative of general adult drinking problems.

Finally, as with most mental health screening inventories, the SMAST and BMAST may be susceptible to responder bias toward social approval. Therefore, counselors may consider coupling these inventories with social desirability measures to increase the veracity of and confidence in the results. For example, counselors should consider administering the Marlowe–Crowne Social Desirability Scale (Crowne & Marlowe, 1960), a free-access instrument that measures social desirability independent of psychopathology, in conjunction with the MAST, BMAST, or SMAST. This could help identify clients who are trying to generate a more positive clinical impression by underreporting drinking problems and who would otherwise screen positive for alcohol use disorders. Enacting these considerations would improve both clinical and research practice with the SMAST and BMAST while preserving the practical advantage of selecting these shorter and quicker inventories to assess problem drinking.

References

*Indicates SMAST-accepted articles included in the meta-analysis.
**Indicates BMAST-accepted articles included in the meta-analysis.
***Indicates BMAST- and SMAST-accepted articles included in the meta-analysis.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: Author.

**Ashman, T. A., Schwartz, M. E., Cantor, J. B., Hibbard, M. R., & Gordon, W. A. (2004). Screening for substance abuse in individuals with traumatic brain injury. *Brain Injury, 18,* 191–202. doi:10.1080/0269905031000149506

*Barry, K. L., & Fleming, M. F. (1990). Computerized administration of alcoholism screening tests in a primary care setting. *The Journal of the American Board of Family Practice, 3,* 93–98.

*Barry, K. L., & Fleming, M. F. (1993). The Alcohol Use Disorders Identification Test (AUDIT) and the SMAST-13: Predictive validity in a rural primary care sample. *Alcohol and Alcoholism, 28,* 33–42.

**Bernadt, M. W., Mumford, J., & Murray, R. M. (1984). Can accurate drinking histories be obtained from psychiatric patients by a nurse conducting screening interviews? *British Journal of Addiction, 79,* 201–206. doi:10.1111/j.1360-0443.1984.tb03856.x

**Bernadt, M. W., Mumford, J., & Taylor, L. (1982). Comparison of questionnaire and laboratory tests in the detection of excessive drinking and alcoholism. *Lancet, 1,* 325–328. doi:10.1016/S0140-6736(82)91579-3

**Bliss, M., Ogley-Oliver, E., Jackson, E., Harp, S., & Kaslow, N. (2008). African American women's readiness to change abusive relationships. *Journal of Family Violence, 23,* 161–171. doi:10.1007/s10896-007-9138-3

Blow, F. C., Gillespie, B. W., Barry, K. L., Mudd, S. A., & Hill, E. M. (1998). Brief screening for alcohol problems in elderly populations using the Short Michigan Alcohol Screening Test–Geriatric Version (SMAST-G). *Alcoholism: Clinical and Experimental Research, 22*(Suppl.), 131A.

*Breakey, W. R., Calabrese, L., Rosenblatt, A., & Crum, R. M. (1998). Detecting alcohol use disorders in the severely mentally ill. *Community Mental Health Journal, 34,* 165–174. doi:10.1023/A:1018793002740

*Buchanan, S. J. (2008). Risk markers associated with physical and psychological abuse by intimate partners against women in substance abuse treatment. *Dissertation Abstracts International: Section B. Sciences and Engineering, 69*(2-B), 1317.

**Chan, A. W. K., Pristach, E. A., Welte, J. W., & Russell, M. (1993). Use of the TWEAK test in screening for alcoholism/heavy drinking in three populations. *Alcoholism: Clinical and Experimental Research, 17,* 1188–1192. doi:10.1111/j.1530-0277.1993.tb05226.x

*Chang, G., Goetz, M. A., Wilkins-Haug, L., & Berman, S. (1999). Identifying prenatal alcohol use: Screening instruments versus clinical predictors. *The American Journal on Addictions, 8,* 87–93. doi:10.1080/105504999305884

*Chang, G., Wilkins-Haug, L., Berman, S., Goetz, M. A., Behr, H., & Hiley, A. (1998). Alcohol use and pregnancy: Improving identification. *Obstetrics and Gynecology, 91,* 892–898.

**Cherpitel, C. J. (1998). Differences in performance of screening instruments for problem drinking among Blacks, Whites and Hispanics in an emergency room population. *Journal of Studies on Alcohol, 59,* 420–426.

*Cleary, P. D., Miller, M., Bush, B. T., Warburg, M. M., Delbanco, T. L., & Aronson. M. D. (1988). Prevalence and recognition of alcohol abuse in a primary care population. *The American Journal of Medicine, 85,* 466–471. doi:10.1016/S0002-9343(88)80079-2

**Connor, J. P., Grier, M., Feeney, G. F., & Young, R. M. (2007). The validity of the Brief Michigan Alcohol Screening Test (bMAST) as a problem drinking severity measure. *Journal of Studies on Alcohol and Drugs, 68,* 771–779. doi:10.15288/jsad.2007.68.771

Corvo, K. (2006). Violence, separation, and loss in the families of origin of domestically violent men. *Journal of Family Violence, 21,* 117–125. doi:10.1007/s10896-005-9011-1

**Cremonte, M., & Cherpitel, C. J. (2008). Performance of screening instruments for alcohol use disorders in emergency department patients in Argentina. *Substance Use & Misuse, 43,* 125–138. doi:10.1080/10826080701212337

*Crews, T. M., & Sher, K. J. (1992). Using adapted short MASTs for assessing parental alcoholism: Reliability and validity. *Alcoholism: Clinical and Experimental Research, 16,* 576–584. doi:10.1111/j.1530-0277.1992.tb01420.x

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24,* 349–354. doi:10.1037/h0047358

*Curtin, J. J., Barnett, N. P., Colby, S. M., Rohsenow, D. J., & Monti, P. M. (2005). Cue reactivity in adolescents: Measurement of separate approach and avoidance reactions. *Journal of Studies on Alcohol, 66,* 332–343. doi:10.15288/jsa.2005.66.332

*Curtis, J. L., Millman, E. J., Joseph, M., Charles, J., & Bajwa, W. K. (1986). Prevalence rates for alcoholism, associated depression and dementia on the Harlem Hospital Medicine and Surgery Services. *Advances in Alcohol & Substance Abuse, 6,* 45–64. doi:10.1300/J251v06n01_04

Erford, B. T. (2013). *Assessment for counselors* (2nd ed.). Belmont, CA: Brooks/Cole.

Erford, B. T., Savin-Murphy, J. A., & Butler, C. (2010). Conducting a meta-analysis of counseling outcome research: Twelve steps and practical procedures. *Counseling Outcome Research and Evaluation, 1,* 19–43. doi:10.1177/2150137809356682

*Fleming, M. F., & Barry, K. L. (1989). A study examining the psychometric properties of the SMAST-13. *Journal of Substance Abuse, 1,* 173–182. doi:10.1016/S0899-3289(88)80020-8

*Fleming, M. F., & Barry, K. L. (1991). The effectiveness of alcoholism screening in an ambulatory care setting. *Journal of Studies on Alcohol, 52,* 33–36. doi:10.15288/jsa.1991.52.33

*Guerriero, J. P. (2015). Impulsiveness among alcohol users. *Dissertation Abstracts International: Section B. Sciences and Engineering, 76*(4-B)(E).

*Hays, R. D., & Revetto, J. P. (1992). Old and new MMPI-derived scales and the Short-MAST as screening tools for alcohol disorder. *Alcohol and Alcoholism, 27,* 685–695.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.

*Hequembourg, A. L., Parks, K. A., & Vetter, C. (2008). Sexual identity and gender differences in substance use and violence: An exploratory study. *Journal of LGBT Issues in Counseling, 2,* 174–198. doi:10.1080/15538600802119301

**Isacsson, S.-O., Hanson, B. S., Janzon, L., Lindell, S.-E., & Steen, B. (1987). Methods to assess alcohol consumption in 68-year-old men: Results from the population study 'Men born in 1914' Malmo, Sweden. *British Journal of Addiction, 82,* 1235–1244. doi:10.1111/j.1360-0443.1987.tb00421.x

*James, L. M., & Taylor, J. (2007). Impulsivity and negative emotionality associated with substance use problems and Cluster B personality in college students. *Addictive Behaviors, 32,* 714–727. doi:10.1016/j.addbeh.2006.06.012

Johnson-Greene, D., McCaul, M. E., & Roger, P. (2009). Screening for hazardous drinking using the Michigan Alcohol Screening Test–Geriatric Version (MAST-G) in elderly persons with acute cerebrovascular accidents. *Alcoholism, Clinical and Experimental Research, 33,* 1555–1561. doi:10.1111/j.1530-0277.2009.00987.x

*Josephson, B. R. (1999). Differences in functional impairment between adults with and without ADHD. *Dissertation Abstracts International: Section B. Sciences and Engineering, 59*(11-B), 6069.

*Landsman, K. J. (2000). Personality, dissociation, and anger in male batterers. *Dissertation Abstracts International: Section B. Sciences and Engineering, 60*(8-B), 4232.

***Lee, D. J., & DeFrank, R. S. (1988). Interrelationships among self-reported alcohol intake, physiological indices and alcoholism screening measures. *Journal of Studies on Alcoholism and Drugs, 49,* 532–537. doi:10.15288/jsa.1988.49.532

**Lehavot, K., & Simoni, J. M. (2011). The impact of minority stress on mental health and substance use among sexual minority women. *Journal of Consulting and Clinical Psychology, 79,* 159–170. doi:10.1037/a0022839

**Lefkowitz, P., & Suljaga-Petchel, K. (1985). The prevalence of alcoholism in acute care general hospital patients. *Mount Sinai Journal of Medicine, 52,* 291–296.

**Lockhart, S. P., Carter, Y. H., Straffen, A. M., Pang, K. K., McLouglin, J., & Baron, J. H. (1986). Detecting alcohol consumption as a cause of emergency general medical admissions. *Journal of the Royal Society of Medicine, 79,* 132–136.

**Luckie, L. F., White, R. E., Miller, W. R., Icenogle, M. V., & Lasoski, M. C. (1995). Prevalence estimates of alcohol problems in a Veterans Administration outpatient population: AUDIT vs. MAST. *Journal of Clinical Psychology, 51*, 422–425. doi:10.1002/1097-4679(199505)51:3<422::AID-JCLP2270510317>3.0.CO;2-B

Magruder-Habib, K., Harris, K. E., & Fraker, G. G. (1982). Validation of the Veterans Alcoholism Screening Test. *Journal of Studies on Alcohol and Drugs, 43,* 910–926. doi:10.15288/jsa.1982.43.910

Magruder-Habib, K., Stevens, H. A., & Ailing, W. C. (1993). Relative performance of the MAST, VAST, and CAGE versus *DSM-III-R* criteria for alcohol dependence. *Journal of Clinical Epidemiology, 46,* 435–441. doi:10.1016/0895-4356(93)90020-2

*Mercer-Lynn, K. B., Flora, D. B., Fahlman, S. A., & Eastwood, J. D. (2013). The measurement of boredom: Differences between existing self-report scales. *Assessment, 20,* 585–596. doi:10.1177/1073191111408229

*Merrill, K. A. (2000). Effects of anxiety and anxiety sensitivity on alcohol use in females. *Dissertation Abstracts International: Section B. Sciences and Engineering,* 3852 (January 2001).

Minnich, A., Erford, B. T., Bardhoshi, G., & Atalay. Z. (2018). Systematic review of the Michigan Alcoholism Screening Test. *Journal of Counseling & Development, 96,* 335–344. doi:10.1002/jcad.12207

*Møller, T., & Linaker, O. M. (2010). Using brief self-reports and clinician scales to screen for substance use disorders in psychotic patients. *Nordic Journal of Psychiatry, 64,* 130–135. doi:10.3109/08039480903274423

*Moore, R. A. (1972). The diagnosis of alcoholism in a psychiatric hospital: A trial of the Michigan Alcoholism Screening Test (MAST). *American Journal of Psychiatry, 128,* 1565–1569. doi:10.1176/ajp.128.12.1565

*Nilssen, O., Ries, R. K., Rivara, F. P., Gurney, J. G., & Jurkovich, G. J. (1994). The CAGE questionnaire and the Short Michigan Alcohol Screening Test in trauma patients: Comparison of their correlations with biological alcohol markers. *The Journal of Trauma, 36,* 784–788.

*Pinchevsky, G. M., Wright, E. M., & Fagan, A. A. (2013). Gender differences in the effects of exposure to violence on adolescent substance use. *Violence and Victims, 28,* 122–144. doi:10.1891/0886-6708.28.1.122

**Pokorny, A. D., Miller, B. A., & Kaplan, H. B. (1972). The Brief MAST: A shortened version of the Michigan Alcoholism Screening Test. *American Journal of Psychiatry, 129,* 342–345. doi:10.1176/ajp.129.3.342

*Reslan, S., & Hanks, R. A. (2014). Factors associated with alcohol-related problems following moderate to severe traumatic brain injury. *Rehabilitation Psychology, 59,* 453–458. doi:10.1037/a0037186

*Robin, R. W., Saremi, A., Albaugh, B., Hanson, R. L., Williams, D., & Goldman, D. (2004). Validity of the SMAST in two American Indian tribal populations. *Substance Use & Misuse, 39,* 601–624. doi:10.1081/JA-120030062

**Rospenda, K. M., Minich, L. M., Milner, L. A., & Richman, J. A. (2010). Caregiver burden and alcohol use in a community sample. *Journal of Addictive Diseases, 29,* 314–324. doi:10.1080/10550887.2010.489450

*Rydon, P., Redman, S. Y., Sanson-Fisher, R. W., & Reid, A. L. A. (1992). Detection of alcohol-related problems in general practice. *Journal of Studies on Alcohol and Drugs, 53,* 197–202. doi:10.15288/jsa.1992.53.197

*Saunders, W. M., & Kershaw, P. W. (1980). Screening tests for alcoholism: Findings from a community study. *British Journal of Addictions, 75,* 37–41. doi:10.1111/j.1360-0443.1980.tb00192.x

Selzer, M. L. (1971). The Michigan Alcoholism Screening Test: The quest for a new diagnostic instrument. *American Journal of Psychiatry, 127,* 1653–1658. doi:10.1176/ajp.127.12.1653

*Selzer, M. L., Vinokur, A., & Van Rooijen, L. (1975). A self-administered Short Michigan Alcoholism Screening Test (SMAST). *Journal of Studies on Alcohol, 36,* 117–126. doi:10.15288/jsa.1975.36.117

*Sheridan, D. P., Johnson, P., Michels, P. J., Marion, L. N., & Fuller, S. G. (1995). Substance abuse in freshmen medical students over time. *Substance Abuse, 16,* 93–98. doi:10.1080/08897079509444709

Shields, A. L., Howell, R. T., Potter, J. S., & Weiss, R. D. (2007). The Michigan Alcoholism Screening Test and its shortened form: A meta-analytic inquiry into score reliability. *Substance Use and Misuse, 42,* 1783–1800. doi:10.1080/10826080701212295

*Silverman, D. C., O'Neill, S., Cleary, P. D., Barwick, C., & Joseph, R. (1992). Recognition of alcohol abuse in psychiatric outpatients and its effect on treatment. *Hospital and Community Psychiatry, 43,* 644–646.

*Simons, J. S., Carey, K. B., & Wills, T. A. (2009). Alcohol abuse and dependence symptoms: A multidimensional model of common and specific etiology. *Psychology of Addictive Behaviors, 23,* 415-427. doi:10.1037/a0016003

Skinner, H. A., & Allen, B. A. (1983). Alcohol dependence syndrome: Measurement and validation. *Journal of Abnormal Psychology, 91,* 199–209. doi:10.1037/0021-843X.91.3.199

**Smith, D. S., Collins, M., Kreisberg, J. P., Volpicelli, J. R., & Alterman, A. I. (1987). Screening for problem drinking in college freshmen. *Journal of American College Health, 36,* 89–94. doi:10.1080/07448481.1987.9938996

*Stein, M. B., Hanna, C., Vaerum, V., & Koverola, C. (1999). Memory functioning in adult women traumatized by childhood sexual abuse. *Journal of Traumatic Stress, 12,* 527–534. doi:10.1023/A:1024775222098

**Tabisz, E., Badger, M., Meatherall, R., Jacyk, W. R., Fuchs, D., & Grymonpre, R. (1991). Identification of chemical abuse in the elderly admitted to emergency. *Clinical Gerontologist, 11,* 27–38. doi:10.1300/J018v11n02_04

*Tangney, J. P., Baumeister, R. F., & Boone, A. L. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality, 72,* 271–324. doi:10.1111/j.0022-3506.2004.00263.x

*Taylor, J., Reeves, M., James, L., & Bobadilla, L. (2006). Disinhibitory trait profile and its relation to Cluster B personality disorder features and substance use problems. *European Journal of Personality, 20,* 271–284. doi:10.1002/per.585

**Underwood, R. A. (2002). Predictors of alcohol use in college students with learning disabilities and attention deficit hyperactivity disorder. *Dissertation Abstracts International: Section B. Sciences and Engineering, 63*(5-B), 2610.

*Vanclay, F., Raphael, B., Dunne, M., Whitfield, J., Lewin, T., & Singh, B. (1991). A community screening test for high alcohol consumption using biochemical and hematological measures. *Alcohol and Alcoholism, 26,* 337–346.

*Westby, M. D., & Ferraro, F. R. (1999). Frontal lobe deficits in domestic violence offenders. *Genetic, Social, and General Psychology Monographs, 125,* 71–102.

***Willenbring, M. L., Christensen, K. J., Spring, W. D., & Rasmussen, R. (1987). Alcoholism screening in the elderly. *Journal of the American Geriatrics Society, 35,* 864–869. doi:10.1111/j.1532-5415.1987.tb02339.x

*Zung, B. J. (1978). Factor structure of the Michigan Alcohol Screening Test. *Journal on Studies of Alcohol, 39,* 56–67. doi:10.15288/jsa.1978.39.56

**Zung, B. J. (1979). Psychometric properties of the MAST and two briefer versions. *Journal of Studies on Alcohol, 40,* 845–859. doi:10.15288/jsa.1979.40.845