ABSTRACT

**Aim** Through a complex combination of direct (face-valid) and indirect (subtle) subscales, the Substance Abuse Subtle Screening Inventory (SASSI) is purported to detect substance use disorders with a high degree of validity regardless of respondent honesty or motivation. This review evaluates empirical evidence regarding the reliability and validity of this widely used screening instrument. **Methods** Source documents were 36 peer-reviewed reports yielding data regarding the SASSI’s internal consistency, test–retest reliability, psychometric structure, convergent and divergent validity and criterion (predictive) validity. **Results** The total N of the studies reviewed equaled 22,110. Internal consistency is high for the overall SASSI and for its direct but not its indirect (subtle) subscales, suggesting that the instrument taps a single face-valid construct. SASSI classifications converged with those from other direct screening instruments, and were also correlated with ethnicity, general distress and social deviance. Studies found test–retest reliability lower than that reported in the test manuals. Sensitivity was found to be similar to that for public domain screening instruments, but on specificity the SASSI appears to yield a high rate of false positives. **Conclusion** No empirical evidence was found for the SASSI’s claimed unique advantage in detecting substance use disorders through its indirect (subtle) scales to circumvent respondent denial or dishonesty. Recommendations for screening and for future research with the SASSI are offered. **Keywords** Adolescent, adult, SASSI, screening, substance abuse.
specifically for its purported ability to detect SUDs regardless of respondents’ honesty or motivation. It contains both direct and indirect scales, which the authors claim operate dynamically together to accurately screen for SUDs in adults and adolescents despite dishonest responding [3–5]. US addiction counselors have reported preferring the SASSI to other screening instruments (i.e. Michigan Alcohol Screening Test (MAST) [6], Addiction Severity Index (ASI) [7]) [8], apparently for its purported ability to circumvent denial. Horrigan & Katz [9, p. 283] asserted that ‘The predictive validity of the SASSI is sufficiently great that it is now the most frequently used empirically based screen in drug and alcohol treatment centers’.

Description of the SASSI

The current adult version (SASSI-3) is comprised of 93 questions on two sides of a single sheet. On the front side are 67 dichotomous (true/false) items with both indirect (i.e. ‘I am rarely at a loss for words’) and direct content (i.e. ‘I have used alcohol or “pot” too much or too often’). The SASSI authors [3] recommend administering the indirect or ‘subtle’ front side first, to prevent biasing by the obvious nature of the back side, which contains 26 Likert-scaled questions directly querying substance use and negative consequences. The adolescent SASSI-A2 is similar: the front side is comprised of 72 items, including 12 items that obviously pertain to substance use in general and eight items that directly query the adolescent’s own substance use. The back side contains an additional 28 direct substance use items.

Scoring and interpretation of the SASSI

There are nine subscale scores on the SASSI-3 and 12 for the SASSI-A2, with separate interpretation criteria for males versus females. Adults exceeding the specified cut-off for any one of nine different decision rules are designated as ‘high probability of having a substance dependence disorder’ [3]. If all nine criteria and the defensiveness score are below cut-offs, the respondent is stated to have a ‘low probability of having a substance dependence disorder’. There is no intermediate category. An elevated defensiveness score is said to signal a possible false negative.

Whereas the adult form was designed to screen for substance dependence, the adolescent version is described as screening for either abuse or dependence. The SASSI-A2 has eight subscales in common with the adult form as well as four unique subscales (see Table 1). Exceeding cut-off on any one of nine decision rules designates the adolescent as ‘high probability of having a substance abuse or substance dependence disorder’ [4, p. 10] and a cut-off value on one [secondary classification scale (SCS)] subscale is used to differentiate abuse from dependence. If none of the nine criteria is met, the adolescent is judged to have a ‘low probability of a substance abuse or substance dependence disorder’ unless either of two validity scores exceeds a cut-off, in which case further assessment is advised.

Development of the SASSI

The original SASSI was released in 1985, and was modified in several ways to develop the 1994 version, SASSI-2 [3]. Two new subscales were added [correctional (COR) and random answering pattern (RAP); see Table 1], the ALD scale (to differentiate alcohol from drug abusers) was deleted, and a defensiveness scale (DEF2, to discriminate substance dependent from non-dependent defensive responders) was reformulated and renamed ‘supplemental addiction measure’ (SAM). Seven items not used in scoring were also added for research purposes.

The SASSI-3 was developed primarily in an attempt to reduce the SASSI-2 false positive rate of 15.5% [3]. The manual [3] describes a primary clinical sample of 2015 respondents, 97% of which ‘were provided by clinicians working in service settings throughout the United States . . . The remaining 3% of the cases (n = 57) were prisoners in a correctional facility or research subjects recruited because they had a family history of alcohol abuse’ [3, p. 23]. A subset of 839 cases in which a diagnosis was available served as the development and cross-validation samples.

The SASSI-A (adolescent version) was first released in 1990 [4]. In 1998, the authors distributed a research version containing the original items plus 45 experimental items to 48 treatment and correctional programs and five schools. From these they received 2326 completed instruments, 53.5% of which contained a Diagnostic and Statistical Manual version IV (DSM-IV) [10] diagnosis of substance use disorders. This subsample of adolescents with SUD diagnoses (n = 1244) was drawn from in-patient psychiatric, addiction treatment, out-patient behavioral health and juvenile corrections programs.

Using these data, the authors developed the SASSI-A2 by adding three direct scales [attitudes (ATT), family–friends risk (FRISK) and symptoms (SYM); see Table 1]; the SCS scale to differentiate substance abuse from dependence, retained 32 non-scored experimental items, and replaced the RAP with a new validity scale [validity check (VAL)]. The original sources of the items used to develop the adolescent and adult SASSI are unclear from the test manuals, as are the procedures used to derive the subscales and decision rules.

Claims for the SASSI

The SASSI-3 and SASSI-A2 manuals claim high rates of accuracy (94%, 94%), sensitivity (94%, 95%), specificity
(94%, 89%), positive predictive power (98%, 98%) and negative predictive power (80%, 75%), with low false positive (6%, 11%) and false negative rates (6%, 5%) for the current adult and adolescent versions, respectively. (A guide to the calculation of these various psychometric indices is found in Fig. 1.)

The authors claimed that the accuracy of the SASSI is unaffected by age, ethnicity, education, comorbidity, adaptive functioning, institutional setting, legal or employment status or by the level of honesty or defensiveness of the respondent, and that the measure detects specific propensity for SUDs rather than general maladjustment [3,4].

The purpose of this review was to evaluate empirical evidence regarding the SASSI. Data from peer-reviewed research were used to examine the psychometric properties of both the adult and the adolescent versions. Specific psychometric indices examined included internal consistency, test–retest reliability, instrument structure, convergent validity, divergent validity and criterion (predictive) validity.

Table 1 Scales of the SASSI.

<table>
<thead>
<tr>
<th>SASSI scales</th>
<th>Abbreviation</th>
<th>Stated purpose [52, p. 31]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct scales</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Face-valid alcohol</td>
<td>FVA</td>
<td>‘Acknowledged use of alcohol’</td>
</tr>
<tr>
<td>Face-valid other drug</td>
<td>FVOD</td>
<td>‘Acknowledged use of other drugs’</td>
</tr>
<tr>
<td>Symptoms</td>
<td>SYM</td>
<td>‘Causes, consequences and correlates of substance misuse’</td>
</tr>
<tr>
<td>Indirect scales</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obvious attributes</td>
<td>OAT</td>
<td>‘Characteristics commonly associated with substance misuse’</td>
</tr>
<tr>
<td>Subtle attributes</td>
<td>SAT</td>
<td>‘Basic personal style similar to substance dependent people’</td>
</tr>
<tr>
<td>Defensiveness</td>
<td>DEF</td>
<td>‘Defensiveness that may or may not be related to substance misuse and that may reflect either an enduring character trait or a temporary reaction to a current situation’</td>
</tr>
<tr>
<td>Supplemental addiction measure</td>
<td>SAM</td>
<td>‘Supplements other scales in some decision rules [but is] not used in clinical interpretation’</td>
</tr>
<tr>
<td>Response set scales</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Random answering pattern*</td>
<td>RAP</td>
<td>‘Assesses whether or not responses are meaningful’</td>
</tr>
<tr>
<td>Family versus control subjects*</td>
<td>FAM</td>
<td>‘Similarity to family members of people who misuse substances’</td>
</tr>
<tr>
<td>Correctional</td>
<td>COR</td>
<td>‘Similarity to people with extensive legal difficulties’</td>
</tr>
</tbody>
</table>

Adolescent version

<table>
<thead>
<tr>
<th>Stated purpose [53]</th>
<th>FRISK</th>
<th>‘Asks directly about the ‘extent to which the client is part of a family/social system that is likely to enable substance misuse’ (p. 32)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attitudes</td>
<td>ATT</td>
<td>‘Asks directly about the ‘client’s attitudes and beliefs regarding substance use’ (p. 33)</td>
</tr>
<tr>
<td>Validity check</td>
<td>VAL</td>
<td>‘Comprised of items that differentiated between SASSI-A2 profiles in which the Decision Rule accurately identified the presence or absence of clinically diagnosed substance use disorders and profiles in which the Decision Rule did not correspond to clinicians’ assessments of the presence or absence of a substance use disorder’ (p. 30)</td>
</tr>
<tr>
<td>Secondary classification scale</td>
<td>SCS</td>
<td>‘Composed of items that differentiated between: (1) People who did not have a substance use disorder, (2) People who have a substance abuse disorder, (3) People who have a substance dependence disorder’ (p. 30)</td>
</tr>
</tbody>
</table>

*Adult measure only.

Figure 1 Guide to calculating psychometric indices
METHOD

Source documents for this review were identified through two literature searches (EbscoHost and Sci Search) during July 2005 and January 2006. Broad search terms (i.e. ‘SASSI’, ‘subtle screening’, ‘substance abuse’) were employed to gather ascertainable peer-reviewed articles citing any use of the SASSI, published between the first release of the SASSI (1985) and the last search date (January 2006). The 36 located studies and two SASSI manuals spanned a wide variety of samples. The adult studies included college students [11–17], pregnant women [9,18–21], mothers of young children [22], community residents [3,13], members of Alcoholics Anonymous [23–26], couples [27], people with traumatic brain injury [28–30] and adults in addiction treatment [3,13,25,26,31], psychiatric settings [3,13,16], correctional programs [3,13,32–34] and welfare or child protective services [9,19,35]. The adolescent studies comprised samples from community or high school [23,36–40], substance abuse [4,23,38,41,42], psychiatric in-patient [4,38,42–44] and correctional settings [4,34,38,42,44–46].

RESULTS

Internal consistency

Internal consistency of a scale is measured through Cronbach’s coefficient alpha (α). The SASSI authors cautioned that because the SASSI scales were not designed to be unidimensional, internal consistency is not a ‘primary consideration’ for this measure [3, p. 26]. Nevertheless, they reported α ranging from 0.27 to 0.95, with the highest values being for the direct scales [face-valid other drug (FVOD), face-valid alcohol (FVA) and SYM]. They also reported α = 0.93 for the entire SASSI [3, p. 26], suggesting redundant measurement of a single construct.

Data from the adult studies [14,15,17,32,34] support high internal consistency for the direct scales (see Table 2). For the direct sales, no study reported alpha coefficients as high as those reported in the test manual [3]. Data from the adult studies revealed generally lower internal consistency for the SASSI subtle scales, with high variability across samples.

No adolescent studies reported internal consistency values for the full measure. However, with an abbreviated face-valid version of the adolescent scale, two studies reported α of 0.74 and 0.66 for their seven- and six-item scales, respectively [37,40].

Instrument structure

The adult SASSI contains 10 scales, three direct and seven indirect. When Gray [34] subjected the responses of 888 adult participants to a confirmatory factor analysis, he found that a 10-factor solution yielded a poor fit. A subsequent exploratory factor analysis yielded a two-factor solution accounting for 53% of the variance in the adult and 36% of the variance in the adolescent sample. These two factors were comprised of items related to (1) alcohol use and (2) drug use, drawn primarily from the direct FVA and FVOD scales. Five indirect scale items loaded with comparable weight upon these two factors, three of which were face-valid items querying alcohol and other substance use.

Two exploratory factor analyses of the adolescent SASSI-A have been reported. In one, a principal components analysis yielded factors labeled as negative affect and fatigue (8%), socialization and conformity (7%), substance abuse (6%) and a lie scale (5%), together accounting for 25% of variance [44]. The second reported a five-factor structure of the SASSI-A, with similar factors of emotional instability, antisocial behavior/conduct and random answering, accounting for 32.5% of the variance [46]. The difference was that Sweet & Saules [46] found two separate factors for substance abuse, drawing on SASSI direct scales: one for alcohol (FVA) and one for drug abuse (FVOD). Like Gray’s finding with the SASSI-2 [34], they also found that the indirect scales did not produce coherent factors, and were highly correlated with the direct scales (cf. [11]).

Convergent validity

Screening instruments for the same disorder are expected to show a high degree of convergence. However, Miller & Lazowski [3,4] suggest that determinations made by the SASSI may not converge with other measures due to the unique contributions of the SASSI’s indirect scales. When compared with the MacAndrew scale, an indirect scale evaluating substance use derived from the MMPI, the SASSI decision rule evidenced modest to moderate correlations (r = 0.53, kappa = 0.34, 0.29) [14,17,32]. In contrast to the decision rule, the SASSI subscales showed more modest (r range = 0.25–0.35) but significant correlations with other MMPI scales [41]. SASSI classifications have been found to converge with those from other direct screening instruments. Converge with the four-item Cut-down, Annoyed, Guilt, Eye-opener (CAGE) is high, with kappa values of 0.61 [14] and 0.49 [17], and a correlation of r = 0.44, improved to 0.58 with a modified CAGE [14]. High convergence was also found with the MAST: r = 0.53 [14], kappa = 0.52 [17], and one low kappa of 0.22 [32]. One correlation of r = 0.43 was reported with the Rutgers Alcohol Problem Index [14].

Divergent validity

Whereas an ideal metric ought to converge with other measures of the same construct, it should not correlate
Table 2 Coefficient alpha (internal consistency) of the SASSI.

<table>
<thead>
<tr>
<th></th>
<th>Adult manual (1)</th>
<th>Adolescent manual (2)</th>
<th>Study 3</th>
<th>Study 4</th>
<th>Study 5 (Sample 1)</th>
<th>Study 6</th>
<th>Study 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>1821</td>
<td>2145</td>
<td>46</td>
<td>57</td>
<td>876</td>
<td>248</td>
<td>230</td>
</tr>
<tr>
<td>Nature of sample</td>
<td>Provided by clinicians</td>
<td>Various adolescent programs</td>
<td>DWI offenders</td>
<td>College students</td>
<td>Justice settings</td>
<td>College students</td>
<td>College students</td>
</tr>
<tr>
<td>Mean age</td>
<td>33</td>
<td>15</td>
<td>31</td>
<td>19</td>
<td>31</td>
<td>24</td>
<td>28</td>
</tr>
<tr>
<td>Version</td>
<td>SASSI-3</td>
<td>SASSI-A2</td>
<td>SASSI-2</td>
<td>SASSI-2</td>
<td>SASSI-2</td>
<td>SASSI-3</td>
<td>SASSI-3</td>
</tr>
<tr>
<td>Coefficient alpha (α)</td>
<td>FVA = 0.93</td>
<td>FVA = 0.91</td>
<td>FVA = 0.91</td>
<td>FVA = 0.89</td>
<td>FVA = 0.90</td>
<td>FVA = 0.86</td>
<td>FVA = 0.92</td>
</tr>
<tr>
<td></td>
<td>FVOD = 0.95</td>
<td>FVOD = 0.93</td>
<td>FVOD = 0.93</td>
<td>FVOD = 0.93</td>
<td>FVOD = 0.94</td>
<td>FVOD = 0.92</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SYM = 0.79</td>
<td>SYM = 0.82</td>
<td>SYM = 0.75</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>OAT = 0.65</td>
<td>OAT = 0.72</td>
<td>OAT = 0.67</td>
<td>OAT = 0.76</td>
<td>OAT = 0.55</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SAT = 0.27</td>
<td>SAT = 0.63</td>
<td>SAT = 0.45</td>
<td>SAT = 0.49</td>
<td>SAT = 0.25</td>
<td>SAT = 0.08</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DEF = 0.63</td>
<td>DEF = 0.64</td>
<td>DEF = 0.36</td>
<td>DEF = 0.52</td>
<td>DEF = 0.66</td>
<td>DEF = 0.56</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SAM = 0.37</td>
<td>SAM = 0.66</td>
<td>SAM = 0.55</td>
<td>SAM = 0.39</td>
<td>SAM = 0.29</td>
<td>SAM = 0.30</td>
<td></td>
</tr>
<tr>
<td></td>
<td>COR = 0.71</td>
<td>COR = 0.61</td>
<td>COR = 0.71</td>
<td>COR = 0.54</td>
<td>COR = 0.72</td>
<td>COR = 0.65</td>
<td></td>
</tr>
<tr>
<td></td>
<td>FAM = 0.33</td>
<td>FRISK = 0.67</td>
<td>FAM = 0.21</td>
<td>FAM = 0.15</td>
<td>FAM = 0.30</td>
<td>FAM = 0.34</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ATT = 0.76</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

substantially with demographics or measures of other constructs. As noted above, the SASSI manuals assert relative independence of SASSI scores from a host of factors such as age, ethnicity, comorbidity and general functioning.

Studies generally support the assertion that adult SASSI scores are unrelated to age [34,41], but have pointed to significant gender differences. Studies have found significant but directionally inconsistent gender effects on scale scores and in the probability of screening positive [37,41,42]. Inconsistent effects of education have also been reported. An adult study (age range = 18–59, M = 40, SD = 10) found a negative relationship between years of education and SASSI scores [30], whereas an adolescent study (age range = 12–18, M = 15.6, SD = 1.2) found no such relationship [41]. Another adolescent study (age range = 12–18) found that learning disabilities and poverty were associated with higher SASSI scores [36].

Studies have more consistently reported a relationship between SASSI scores and ethnicity. Ethnic minorities have been found to be significantly more likely to be classified by the SASSI decision rules as ‘high probability’ of SUD relative to Caucasians [36,39], and to score higher on the DEF, RAP [44] and COR scales [38].

Independent studies also suggest that SASSI scores and classifications are influenced substantially by general distress and deviance. Positive relationships with SASSI scales have been reported for conduct disorder [38], depression [19,20,38], social anxiety [16], general distress and traumatic histories [19,22], and suicidal ideation or attempts [39,41]. The indirect scales and associated decision rules seem particularly linked to more global distress and psychopathology [16,30,46].

Table 3 Criterion validity of the SASSI.

<table>
<thead>
<tr>
<th>Version</th>
<th>n</th>
<th>Criterion</th>
<th>Sens</th>
<th>Spec</th>
<th>FP</th>
<th>FN</th>
<th>OAR</th>
<th>PPP</th>
<th>NPP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adult manual [3]</td>
<td>839</td>
<td>SASSI-3</td>
<td>DSM diagnosis</td>
<td>94%</td>
<td>94%</td>
<td>1%</td>
<td>5%*</td>
<td>94%</td>
<td>98%</td>
</tr>
<tr>
<td>Areth et al. 2001 [29]</td>
<td>78</td>
<td>SASSI-3</td>
<td>Clinician DSM</td>
<td>85%</td>
<td>63%</td>
<td>15%</td>
<td>9%</td>
<td>76%</td>
<td>76%</td>
</tr>
<tr>
<td>Ashman et al. 2004 [30]</td>
<td>223</td>
<td>SASSI-3</td>
<td>Clinician DSM</td>
<td>72%</td>
<td>82%</td>
<td>12%</td>
<td>10%</td>
<td>78%</td>
<td>68%</td>
</tr>
<tr>
<td>Clemens 2002 [15]</td>
<td>248</td>
<td>SASSI-3</td>
<td>Computer CIDI</td>
<td>65%</td>
<td>89%</td>
<td>10%</td>
<td>5%</td>
<td>85%</td>
<td>51%</td>
</tr>
<tr>
<td>Fuller et al. 1994 [28]</td>
<td>10</td>
<td>SASSI</td>
<td>Clinician DSM</td>
<td>60%</td>
<td>100%</td>
<td>0</td>
<td>20%</td>
<td>80%</td>
<td>100%</td>
</tr>
<tr>
<td>Horrigan &amp; Piazza 1999 [18]</td>
<td>1251</td>
<td>SASSI</td>
<td>Nurse query</td>
<td>59%</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Lindquist et al. 1997 [27]</td>
<td>25</td>
<td>SASSI</td>
<td>Men in recovery</td>
<td>52%</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Piazza 1996 [43]</td>
<td>203</td>
<td>SASSI</td>
<td>Clinician DSM</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>90%</td>
<td>–</td>
</tr>
<tr>
<td>Svanum &amp; McGrew 1995 [12]</td>
<td>495</td>
<td>SASSI</td>
<td>Computer CDIS</td>
<td>33%</td>
<td>87%</td>
<td>12%</td>
<td>8%</td>
<td>81%</td>
<td>25%</td>
</tr>
<tr>
<td>Swartz 1998 [31]</td>
<td>1837</td>
<td>SASSI-2</td>
<td>Clinician</td>
<td>87%</td>
<td>49%</td>
<td>11%</td>
<td>10%</td>
<td>78%</td>
<td>86%</td>
</tr>
<tr>
<td>Adolescent manual [4]</td>
<td>1244</td>
<td>SASSI-A2</td>
<td>DSM diagnosis</td>
<td>95%</td>
<td>89%</td>
<td>2%*</td>
<td>4%*</td>
<td>94%</td>
<td>98%</td>
</tr>
<tr>
<td>Bauman et al. 1999 [38]</td>
<td>79</td>
<td>SASSI-A</td>
<td>Clinician</td>
<td>–</td>
<td>–</td>
<td>35%</td>
<td>3%</td>
<td>62%</td>
<td>–</td>
</tr>
</tbody>
</table>

Sens = sensitivity, Spec = specificity, FP = false positive rate, FN = false negative rate. OAR = overall agreement rate, PPP = positive predictive power, NPP = negative predictive power; *pp. 56; †substance use disorders in remission were considered as criterion positive; ‡pp. 31; †adult measure used with an adolescent sample. The rates of false positives and false negatives reported here are lower than those stated in the test manual because we used as the denominator all cases (see Fig. 1). –, Not provided or deducible from data provided within the article.

Criterion validity of the SASSI

The ‘gold standard’ against which the SASSI has been evaluated is a SUD diagnosis by a clinical or a computer-based interview (DSM-IV-TR [47]). Convergence of SASSI classifications with diagnostic interviews can be summarized in seven indices computed from the $2 \times 2$ classification table illustrated in Fig. 1. Few studies have reported all seven indices, but for some of those shown in Table 3 we were able to derive the $2 \times 2$ classification table, from which all seven can be computed. A few articles reported only the overall agreement rate. We strongly recommend that future reports should include at least the complete $2 \times 2$ classification table.

Sensitivity

Of all cases with a true positive SUD diagnosis, what percentage are detected by the SASSI using its standard scoring rules? For a screening instrument, this sensitivity index is of primary concern. Properly used, a screener simply indicates the need for more careful evaluation for a possible SUD. Thus it is desirable for a screener to have high sensitivity and a low false-negative rate.

The manual for the SASSI-3 [3] claims sensitivity of 94%, a rate that was not replicated in the eight independent evaluations summarized in Table 3. The reported sensitivity rates for the SASSI vary from 33% in a college sample ($n = 495$) [12] to 87% in a corrections sample ($n = 1837$) [33], averaging 69.8% (with studies weighted
by N). That is, the SASSI has detected roughly seven of 10 cases with actual SUD diagnoses, on average. This is comparable to the sensitivity rate of briefer, face-valid screening scales such as the MAST [6] and CAGE [48], and slightly lower than that of the Alcohol Use Disorders Identification Test (AUDIT) [49]. Ashman and colleagues [30] found that the brief MAST had higher accuracy, sensitivity and specificity than the SASSI in detecting lifetime diagnosis of SUD.

What proportion of cases identified as probable SUD by the SASSI are detected by the direct scales alone? The test manual reports that the direct scales (rules 1–3) detected only 79% of actual SUDs, whereas adding the indirect (subtle) scales increased sensitivity to 94% [3, p. 30]. That is, 84% (79/94) of all SUDs cases that were identified correctly by the SASSI were detected by the direct scales alone.

Other studies have found that 87% [33] and 89% [32] of cases classified as probable SUDs were so determined based on the direct scales alone, and only 11–13% uniquely by the indirect scales. Several studies have reported that the direct scales of the SASSI perform as well as or better than the total SASSI with the indirect scales included [14,15,30,32,34,44]. Studies have also reported generally high kappa values for positive/negative screening classifications from the SASSI versus commonly used direct screening tools such as the CAGE and MAST [14,17]. Others have reported that the SASSI did not contribute additional unique sensitivity in predicting actual SUD diagnosis, above that accounted for by direct screeners such as the MAST [12,28].

Pajulo and colleagues [20] found the indirect scales markedly increased the rate of cases screening positive. Of their sample of pregnant women (n = 375), 6.1% screened positive on a Finnish version of the SASSI. Of these 23 women, only two screened positive on the direct scales, and 21 (91.3%) through the indirect scales. Because this study had no criterion measure, however, the accuracy of these classifications is unknown.

A crucial piece of missing information in the literature, reported in no study that we could find, is the positive predictive power for cases classified as ‘probable SUD’ based on SASSI indirect scales alone.

The SASSI has been claimed to be impervious to efforts to ‘fake good’. The items that comprise three of the indirect scales [subtle attributes (SAT), DEF, SAM] were derived from responses of participants told to conceal evidence of SUD and/or to ‘fake good’ [3, pp. 2–3]. Other studies have not found the SASSI to be resistant to false responding. In a sample of DUI offenders who screened positive on the SASSI-2, 71% of those told to ‘fake good’ on a second administration were able to change their classification status to negative [32]. Moreover, 63% of those screened negative (non-substance-dependent) by the SASSI-2 sub-

sequently scored as positive when instructed to simulate a SUD. In another study, after directing their college sample to ‘fake good’, ‘fake bad’ or respond according to the standard instructions, Myerholtz & Rosenberg [14] found that a relatively equivalent percentage of their sample were classified as substance-dependent between the standard (35%) and ‘fake good’ (36%), but the percentage of decision-rule positives jumped to 57% for the sample who had been instructed to ‘fake bad’. When examined, this high percentage stemmed from an increased contribution of positives from the indirect scales.

Is sum, we found no independent empirical evidence that the SASSI is more sensitive or accurate or less susceptible to falsification in screening for SUDs than simpler direct scales available in the public domain. No study has managed either to replicate the high sensitivity rates reported in the SASSI test manuals, or to demonstrate a unique additive contribution to accuracy from the SASSI indirect scales.

Specificity

Specificity (the absence of false positives) is of somewhat less concern when an instrument is truly used as a screener. The SASSI, however, is purported to be more than this, able to detect SUDs regardless of a respondent’s honesty. This could lead users to believe the instrument when it contradicts direct self-report, although above-reviewed research on sensitivity provides no support for this claim. Of greater concern, the SASSI is commonly used to make treatment recommendations. The SASSI manuals provide guidelines for assigning people to various levels of treatment intensity, while acknowledging that there is no scientific basis for these guidelines [3, pp. 13–21]. When treatment decisions, including court mandates, are made on the basis of an instrument, it is no longer a screener, and specificity becomes a serious concern. To what extent can the SASSI incorrectly identify people as ‘probable SUD’?

Of 839 cases reported in the adult manual [3, p. 28], only 10 (1.2%) were classified incorrectly as ‘probable SUD’ (8 out of a total of 835 on the basis of indirect scales alone; p. 31). Of those identified by the full SASSI as probable SUD, only 1.6% were classified incorrectly.

With the exception of one study with 10 brain-injured participants [28], independent studies have failed to replicate the specificity rates reported by the SASSI authors (see Table 3). Excluding the values stated in the SASSI manuals [3,4,13], the N-weighted mean of specificity rates is 62%, which would indicate that on average 38% of those screening positive on the SASSI would be classified incorrectly. Contrary to data reported in the SASSI manuals, the overall false positive rate (N-weighted mean = 12% of all cases) is found consistently to be
higher than the SASSI’s false negative rate (N-weighted mean = 9%). Of the 625 SASSI classification errors reported in the studies in Table 3 (excluding the test manuals), 349 (56%) were false positives.

Why would there be so many false positives? One apparent factor is that the SASSI screens for life-time SUD, and classifies as ‘probable SUD’ people with a past history of SUD who no longer meet diagnostic criteria. Rogers [44], reporting on data collected from 317 adolescent respondents, found that of the 19 adolescents reporting no current use of substances, at least 79% were classified by the SASSI-A as having a high probability of SUD. This has also been a problem with the MAST, which asks ‘ever’ questions and thereby detects people with past but not current diagnoses [50].

Independent studies of the SASSI suggest that it is not as specific to current SUDs as implied by the test manuals. Significant numbers of people are classified incorrectly as having a high probability of SUDs. As noted within the adolescent manual [4, p. 36], such misclassification is increased by use of the indirect (subtle) scales, a finding consistent with prior research on direct and indirect screening scales [1].

Test–retest reliability

Another important question is the extent to which an instrument gives the same answer when administered twice. The SASSI authors reported 2-week test–retest correlations averaging 0.96 for subscales of the adult version [3] and 0.85 for the adolescent scale [4]. Perhaps more importantly, the manuals reported stability of 0.92–1.00 for the adult and 0.81 to 0.92 for the adolescent binary classifications of risk, with only 6% of the adolescent classifications changing over a 2-week period [3,4].

Independent studies have been unable to reproduce results in this range and have found lower reliability for the SASSI. Myerholtz & Rosenberg [14] reported test–retest correlations for subscales averaging $r = 0.66$ at 2 weeks and 0.61 at 4 weeks (range = −0.03 to 0.93). Highest stability was found for the direct scales [FVA, FVOD], with less stability for the subtle (indirect) scales that are alleged to measure stable traits. Further doubt about the trait nature of the subtle scales is raised by findings of significant decreases in mean scores after treatment [25,42].

There are also concerns about the stability of risk classifications. One study [17] found significant shifts in SASSI classifications over a 1-week interval. The 1-week stability (phi) coefficient for SASSI classification was 0.63, compared with 0.86 for a direct scale (the CAGE), 0.51 for an indirect scale (MacAndrew; MAC) and 0.38 for the MAST. The total percentage classified as being at risk also dropped from 30% to 21% in this 1-week period. Myerholtz & Rosenberg [14] found that 9% of cases changed classification without treatment over a 2-week period and 19% over a 4-week period. Classification as ‘high risk’ for SUD decreased after 12-week treatment to such an extent (40%) that Schmidt [25] suggested using SASSI scores to assess recovery.

**DISCUSSION**

Beyond the developmental research reported by the scale’s authors [3,4], we found 36 articles reporting data from the SASSI. The internal consistency of the overall SASSI and of its direct self-report scales of alcohol (FVA) and drug use (FVOD) is quite high ($\alpha > 0.90$), indicating substantial item redundancy in measuring a single construct. For the indirect and validity scales, however, internal consistency tends to be fair to poor, suggesting that these scales do not tap a coherent attribute. This is consistent with the results of factor analyses of the SASSI.

The adult SASSI tends to yield two factors centered on the direct reporting of alcohol (FVA) and drug use (FVOD). The adolescent SASSI shows one or two factors for direct reporting of substance use, and three others reflecting general emotional distress, social non-conformity and random answering. The latter three draw items from multiple subscales.

On convergent validity, SASSI direct scale scores and classifications generally agree with the results of other direct screening instruments such as the CAGE [48] and MAST [6]. SASSI scores (especially indirect scales) and classifications have been found consistently to be related to ethnicity (with minorities more likely to screen positive), general distress and social deviance.

Independent studies indicate that test–retest reliability of the SASSI scales is lower than reported in the manuals. Classifications of individuals as ‘high probability’ versus ‘low probability’ were found to fluctuate significantly over periods of 1–4 weeks.

Beyond the direct self-report of alcohol and drug use, it is unclear what the SASSI is measuring. It is purported to tap traits associated with life-time incidence of SUDs, but the indirect scales show low internal consistency, and their factor structure does not resemble the instrument’s scales. The traits alleged to be predictive of SUDs and to be measured by the SASSI are not specified clearly, but both factor analyses and correlations with other measures indicate that the SASSI assesses global distress and social deviance. This is also consistent with the finding that indirect scale scores decline with treatment, which would not be expected with stable traits. The subtle (indirect) scales contain many face-valid items asking about substance use and consequences.

Another interpretive difficulty is the SASSI’s timeframe. The indirect scales are designed to assess life-time...
SUDs, and on direct scales there is an option to specify a period of use, with life-time use being the apparent default. Because the SASSI is used typically to screen for current disorders, this would be expected to produce an undesirable rate of false positives (as is reported in several studies) for people with past but not current SUDs [33].

In sum, even allowing for the wide differing administrations of both adult and adolescent SASSI in their respective editions, no independent peer-reviewed substantiation was found for the SASSI’s claimed unique advantage in detecting SUDs through its indirect (subtle) scales, by circumventing respondent denial or dishonesty. The SASSI’s rates of false negatives (missed cases) are reportedly similar to those for simpler direct scales that are available in the public domain, such as the CAGE [48], MAST [6] and AUDIT [49]. The AUDIT was specifically developed and has been shown to be impervious to national differences [51], whereas independent studies suggest an overclassification bias for the SASSI when used with ethnic minorities. While showing no sensitivity advantage over direct scales, the SASSI does appear to yield significant rates of false positives.

Misuses of the SASSI

False positives might not be of great concern if the SASSI were used only as a screener, to signal the need for more thorough evaluation. In a number of studies, however, the SASSI has been the sole determinant of SUD diagnoses [11,20–22,24] or prevalence [23,36,37,39,42,45]. In community practice, the SASSI is apparently being used to inform treatment decisions and correctional dispositions, and to override self-report, as to determine ‘a client’s degree of criminal involvement . . . when other data such as arrest history are unavailable’ [33, p. 349]. Criteria for making treatment intensity assignments are specified in the adult and adolescent SASSI manuals [3,4], with the caveat that there is no scientific basis for their use.

Recommendations

We found no empirical evidence to support claims that the indirect scales of the SASSI offer a unique or additive advantage in correctly detecting current substance use disorders. It would therefore be cost-effective to use public domain screening instruments which are available free of charge and perform at least as well as the direct scales of the SASSI. In addition, based on this review of published research on the SASSI, we offer the following recommendations:

1. The SASSI should not be used to inform the making of diagnoses, treatment recommendations or dispositional/correctional decisions with regard to SUDs.

2. In future research, when reporting classifications from the SASSI relative to a criterion measure, the full 2 x 2 classification table should be specified. It would also be useful to report the number of new cases that were classified as ‘high risk’ by each successive decision rule.

3. The SASSI authors be invited to produce, for independent peer-review, externally established evidence for each of the quantified claims of reliability and validity made for SASSI in their own literature. If the authors are unable to do so in any case, the authors should make it clear that the authors’ claims in that case are based merely on the conclusions of the authors or withdraw the uncorroborated claim from their literature.

References