Review article

Reliability of Instruments Measuring At-Risk and Problem Gambling Among Young Individuals: A Systematic Review Covering Years 2009–2015

Robert Edgren a,b,*, Sari Castrén, Ph.D. b,c, Marjukka Mäkelä, M.D., Ph.D., M.Sc. d,e, Pia Pörtfors, M.Sc. f, Hannu Alho, M.D., Ph.D. b,c, and Anne H. Salonen, Ph.D. b,c

a Institute of Behavioural Sciences, Faculty of Behavioural Sciences, University of Helsinki, Helsinki, Finland
b Department of Tobacco, Gambling and Addiction, National Institute for Health and Welfare, Helsinki, Finland
c Institute of Clinical Medicine, Faculty of Medicine, University of Helsinki, Helsinki, Finland
d Finnish Office for HTA (FINOHTA) at National Institute for Health and Welfare, Helsinki, Finland
e Department of General Practice, Institute of Public Health, University of Copenhagen, Copenhagen, Denmark
f Department of Information Services, National Institute for Health and Welfare, Helsinki, Finland

Article history: Received December 29, 2015; Accepted March 8, 2016
Keywords: Gambling; Adolescence; Measurement; Reliability; Validity; Systematic review

ABSTRACT

This review aims to clarify which instruments measuring at-risk and problem gambling (ARPG) among youth are reliable and valid in light of reported estimates of internal consistency, classification accuracy, and psychometric properties. A systematic search was conducted in PubMed, Medline, and PsycInfo covering the years 2009–2015. In total, 50 original research articles fulfilled the inclusion criteria: target age under 29 years, using an instrument designed for youth, and reporting a reliability estimate. Articles were evaluated with the revised Quality Assessment of Diagnostic Accuracy Studies tool. Reliability estimates were reported for five ARPG instruments. Most studies (66%) evaluated the South Oaks Gambling Screen Revised for Adolescents. The Gambling Addictive Behavior Scale for Adolescents was the only novel instrument. In general, the evaluation of instrument reliability was superficial. Despite its rare use, the Canadian Adolescent Gambling Inventory (CAGI) had a strong theoretical and methodological base. The Gambling Addictive Behavior Scale for Adolescents and the CAGI were the only instruments originally developed for youth. All studies, except the CAGI study, were population based. ARPG instruments for youth have not been rigorously evaluated yet. Further research is needed especially concerning instruments designed for clinical use.

IMPLICATIONS AND CONTRIBUTION

Rigorous psychometric evaluation of instruments measuring at-risk and problem gambling in young people has not been accomplished yet. Reporting estimates of internal consistency from previous articles is not enough, considering the weak theoretical foundation of these instruments. Reliability testing in population-based studies and validation for clinical use are needed.

Adolescents have persistently been reported to have higher problem gambling rates than the adult population [1,2]. Furthermore, initiation of gambling at a young age has consistently been identified as a risk factor for developing gambling-related problems [3,4]. Therefore, adequate tools for identifying both at-risk and problem gambling (ARPG) among adolescents are of utmost importance [5]. Here, ARPG refers to a wider spectrum of problematic adolescent gambling [6,7].

Robert Edgren will handle correspondence at all stages of refereeing and publication. Disclaimer: The Ministry of Social Affairs and Health had no role in the study design, collection of the data, analysis or interpretation of the results of the manuscript, writing of the report or any phase of the publication process.

* Address correspondence to: Robert Edgren, Department of Tobacco, Gambling and Addiction, National Institute for Health and Welfare, P.O. Box 30, FI-00271, University of Helsinki, Helsinki, Finland.
E-mail address: robert.edgren@helsinki.fi (R. Edgren).
prevalence rates of adolescent problem gambling in relation to adult prevalence rates have raised concerns about the validity of screening instruments, notably whether adolescent rates are exaggerated [18].

In 2010, Stinchfield [9] conducted a critical review of youth problem gambling assessment instruments, identifying four instruments. Two of these were simply adaptations of adult instruments where item wording was modified to better represent potential adolescent-specific adverse consequences [8,9]. These were the South Oaks Gambling Screen Revised for Adolescents (SOGS-RA) and the Diagnostic Statistical Manual IV (Multiple Response format) adapted for Juveniles (DSM-IV-J; DSM-IV-MR-J; Table 1). The Massachusetts Gambling Screen (MAGS) was developed on a sample of high school students, but according to developers, it could be used for both adolescents and adults [9,17]. The Canadian Adolescent Gambling Inventory (CAGI) was the only instrument purposefully developed for an adolescent population.

Previous studies on adolescent gambling have mainly been nontheoretical, quantitative, prevalence based, descriptive, and school based [1]. ARPG instruments are used to make judgments at the individual level, and for decision-making at a societal level, making it essential for such tools to be valid and

<table>
<thead>
<tr>
<th>Instrument [ref]</th>
<th>Content and structure</th>
<th>Items and time frame</th>
<th>Classification cutoff score</th>
<th>Strengths and weaknesses</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOGS-RA [10]</td>
<td>Signs and symptoms of problem gambling and its negative consequences.</td>
<td>12 Items with two response options (yes/no) scored 0–1. Four additional items provide insight to an individual's gambling, but not used in scoring. Time frame: Past year</td>
<td>Recent studies have preferred the narrow criteria: sum score 2–3 = at-risk gambling; ≥4 = problem gambling</td>
<td>Widespread use.</td>
</tr>
<tr>
<td>DSM-IV-J [15]</td>
<td>Based on the DSM-IV criteria. The DSM-IV-MR-J is a modified version of the DSM-IV-J featuring simpler language, fewer details and multiple response options.</td>
<td>12 Items with two response options (yes/no) scored 0–1. Time frame: Past year</td>
<td>Sum score ≥4 = problem gambling</td>
<td>The DSM-IV-J is considered more conservative than the SOGS-RA. It fails to measure the DSM-IV criterion “loss of control.” The exclusion criterion is claimed to be premature. The multiple-response format of the DSM-IV-MR-J is collapsed when calculating the scores.</td>
</tr>
<tr>
<td>MAGS [15]</td>
<td>Psychological and social problems related to gambling. Developed using items from the Short Michigan Alcoholism Screening Test [110]. DSM-IV was used as a reference standard in the development process.</td>
<td>14 Items; seven items are scored in a scale based on item weights from a discriminant function analysis; yes/no. Time frame: Past year</td>
<td>Each item is scored 0 for no and 1 for yes. Each item score is multiplied by a weight and then summed along with constant using a weighted scoring algorithm derived from a discriminant function analysis. The instrument classifies respondents into nonpathologic gambling, transitional gambling, or pathologic gambling [9,17].</td>
<td></td>
</tr>
<tr>
<td>CAGI [18,19]</td>
<td>Five domains: 1) gambling problem severity, 2) psychological consequences, 3) social consequences, 4) financial consequences, and 5) loss of control.</td>
<td>45 Items, using a four-point multiple response format. Time frame: Past 3 months</td>
<td>Three categories: no problem (score 0–1), low to moderate severity (score 2–5), high severity (score ≥6)</td>
<td>Includes an item inquiring about gambling activity of a fictitious gambling form, to test the validity of self-report. Includes a nine-item subscale (GPSS) consisting of the discriminatively best items from the five domains. Administration time is up to 20 minutes while the SOGS-RA, DSM-IV-J/DSM-IV-MR-J and MAGS have been estimated to take a maximum of 10 minutes [9]. All information up to year 2015 is based on three reports. Concludes to have a good validity and reliability among adolescents in Korea. Information about the instrument is limited.</td>
</tr>
<tr>
<td>GABSA [79]</td>
<td>Four domains: 1) loss of control; 2) life dysfunction from problem gambling; 3) gambling experience; 4) social dysfunction from problem gambling.</td>
<td>25 Items. Time frame: Not specified</td>
<td>Three categories: nongambling, nonproblem gambling, and problem gambling; no cutoff scores for classification specified</td>
<td>Includes an item inquiring about gambling activity of a fictitious gambling form, to test the validity of self-report. Includes a nine-item subscale (GPSS) consisting of the discriminatively best items from the five domains. Administration time is up to 20 minutes while the SOGS-RA, DSM-IV-J/DSM-IV-MR-J and MAGS have been estimated to take a maximum of 10 minutes [9]. All information up to year 2015 is based on three reports. Includes an item inquiring about gambling activity of a fictitious gambling form, to test the validity of self-report. Includes a nine-item subscale (GPSS) consisting of the discriminatively best items from the five domains. Administration time is up to 20 minutes while the SOGS-RA, DSM-IV-J/DSM-IV-MR-J and MAGS have been estimated to take a maximum of 10 minutes [9]. All information up to year 2015 is based on three reports. Concludes to have a good validity and reliability among adolescents in Korea. Information about the instrument is limited.</td>
</tr>
</tbody>
</table>

CAGI = Canadian Adolescent Gambling Inventory; DSM-IV-J/DSM-IV-MR-J = the Diagnostic Statistical Manual IV (Multiple Response format) adapted for Juveniles; GABSA = Gambling Addictive Behavior Scale for Adolescents; GPSS = Gambling Problem Severity Subscale; MAGS = Massachusetts Gambling Screen; SOGS-RA = the South Oaks Gambling Screen–Revised for Adolescents.
Reliable [9]. Overall, there is a lack of validated instruments to assess ARPG among youth [1,9,20]. To our knowledge, a systematic review has not yet been conducted on this subject.

This systematic review aims to summarize existing evidence on the psychometric evaluations of instruments designed to evaluate ARPG among youth. The transition from being an adolescent to a young adult, and simultaneously becoming able to gamble legally, is a critical phase concerning gambling behavior. The age range of youth is limited to a maximum of 28 years [21]. Specifically, we wish to clarify which instruments measuring ARPG among youth are reliable and valid for both population-based and clinical studies in light of reported estimates of internal consistency, classification accuracy, and psychometric properties. Consequently, the aim was to identify suitable instruments presently available and provide insight on what branches of the field require further investigation.

Methods

Search strategy

A structured electronic search was conducted (November 2014) in PubMed, Medline, and PsycInfo databases, covering articles published between 2009 and 2014. The search result was processed and reported according to methods recommended in the PRISMA statement [22,23]. This study complements the earlier reviews [1,9,24,25]. The latest two reviews [1,9] were published in 2010; therefore, to fill in the gap between preparing the articles and publishing them, the year 2009 was included in our search.

The search terms were categorized using the Cochrane handbook guidelines for formulating review questions and inclusion criteria PICO [26]: P (patient, i.e., population), I (intervention, i.e., instrument), C (comparator, i.e., reference instrument), and O (outcome, i.e., reliability). The search strategy for PubMed is in the Supplementary Data; further search strategy details across bibliographic databases are available on request. On completion, the searches from each database were documented and references imported into RefWorks, where duplicates were eliminated.

Reference lists of the included articles and of identified review articles were scrutinized to find articles unrecognized in initial searches, resulting in eight additional articles for the systematic review. Database searches were updated in November 2015. The flow chart of the article selection process is presented in Figure 1.

Eligibility criteria

Original research articles written in English were accepted. Publications in peer-reviewed journals, doctoral theses, and institutional reports were accepted. Both population-based samples and clinical samples were eligible. First, 822 abstracts were evaluated using the following exclusion criteria: (1) non–gambling-related research topic; (2) sample age >28 years; (3) no gambling instrument employed; (4) case study, commentary, editorial, or letter; and (5) other (specification required; for example, papers involving only families of young gamblers or youth whose parents gambled were excluded).

Second, 445 full texts of the articles, doctoral theses [29,38], and reports [53,65,78] (later referred as articles) were evaluated using inclusion criteria matching the PICO:

P: ≤28 years of age; I: instrument designed to assess youth gambling; and O: instrument reliability reported.

A comparator criterion was not required for inclusion but was included for labeling the articles. It referred to whether the study used a reference standard in addition to the primary instrument for gambling assessment. Six articles repeating the same reliability estimates as the primary source were excluded. Altogether, 50 articles were included.

In all phases, two researchers independently assessed all articles. The joint probability of agreement was 90.9% for the exclusion of articles based on the abstracts. The joint probability of agreement for the inclusion of articles based on full-text evaluation was 92.4%. Disagreement in evaluation was resolved by discussion and with a third independent researcher.

Quality assessment

The articles were appraised using the revised Quality Assessment of Diagnostic Accuracy Studies tool (QUADAS-2) [27], which assesses the risk of bias and the applicability of articles. QUADAS-2 questions were tailored for our review, tested on a subsample of articles and modified to ensure unambiguous assessment between researchers. An important change was omitting the evaluation of risk of bias regarding the reference standard. This decision was made because both the index test and reference standard (if measured) were applied identically within the gambling context. Information pertaining to the reference standard was inferred from the applicability assessment. The risk of bias was assessed by five questions (Table 2), and each guiding question required an answer of “high,” “low,” or “unclear” risk of bias, where unclear risk refers to suboptimal reporting. The applicability of articles was evaluated with three questions (Table 2), and each item was scored as either “good,” “poor,” or “unclear” applicability. The joint probability of agreement was 89.8%. Disagreement was resolved by discussion (see Eligibility criteria section).

Data extraction

Information from the articles was compiled into tables (Tables 3 and 4). Data from external articles (e.g., methodological reports) were used when necessary to complete study information.

Results

Of the 50 articles, 33 dealt with the SOGS-RA and 12 with the DSM-IV-J or DSM-IV-MR-J [collectively referred to as DSM-IV-(MR)-J] as the primary ARPG instrument. Three articles evaluated the MAGS, one evaluated the CAGI, and one evaluated the Gambling Addictive Behavior Scale for Adolescents (GABSA). Four studies primarily using the SOGS-RA also used one of the two DSM instruments, and one study primarily using the DSM-IV-MR-J also used the SOCS-RA. All samples were general population based, except the study on the CAGI, which used a sample from a substance use center. About half of the articles (n = 26) used cross-sectional data from local schools, including two pairs of articles based on the same data sets. Eight articles had a longitudinal study design. Five articles had cross-sectional nationally representative samples and another nine articles used...
cross-sectional data from nationally representative school samples, including one pair of articles using partially the same data set. Two studies had local school samples and used experimental designs. Articles originated from 17 different countries, including 16 from the United States, 10 from Canada, 7 from Italy, 3 from Australia, and 2 from Norway and 1 article from China, England,

Figure 1. Flow chart of the retrieval and inclusion/exclusion process for articles used in review.

Table 2
Quality assessment of included articles using the revised Quality Assessment of Diagnostic Accuracy Studies tool

<table>
<thead>
<tr>
<th>Articles</th>
<th>Risk of bias n (%)</th>
<th>Applicability n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patient selection</td>
<td>Index test</td>
</tr>
<tr>
<td>All (n = 50)</td>
<td>Low risk/good applicability</td>
<td>24 (48.0)</td>
</tr>
<tr>
<td></td>
<td>High risk/poor applicability</td>
<td>25 (50.0)</td>
</tr>
<tr>
<td></td>
<td>Unclear</td>
<td>1 (2.0)</td>
</tr>
<tr>
<td>SOGS-RA (n = 33)</td>
<td>Low risk/good applicability</td>
<td>14 (42.4)</td>
</tr>
<tr>
<td></td>
<td>High risk/poor applicability</td>
<td>19 (57.6)</td>
</tr>
<tr>
<td></td>
<td>Unclear</td>
<td>0 (0)</td>
</tr>
<tr>
<td>DSM-IV-J/DSM-IV-MR-J (n = 12)</td>
<td>Low risk/good applicability</td>
<td>5 (41.7)</td>
</tr>
<tr>
<td></td>
<td>High risk/poor applicability</td>
<td>6 (50.0)</td>
</tr>
<tr>
<td></td>
<td>Unclear</td>
<td>1 (8.3)</td>
</tr>
<tr>
<td>MAGS (n = 3)</td>
<td>Low risk/good applicability</td>
<td>3 (100.0)</td>
</tr>
<tr>
<td></td>
<td>High risk/poor applicability</td>
<td>0 (0)</td>
</tr>
<tr>
<td></td>
<td>Unclear</td>
<td>0 (0)</td>
</tr>
<tr>
<td>CAGI (n = 1)</td>
<td>Low risk/good applicability</td>
<td>1 (100.0)</td>
</tr>
<tr>
<td></td>
<td>High risk/poor applicability</td>
<td>0 (0)</td>
</tr>
<tr>
<td></td>
<td>Unclear</td>
<td>0 (0)</td>
</tr>
<tr>
<td>GABSA (n = 1)</td>
<td>Low risk/good applicability</td>
<td>1 (100.0)</td>
</tr>
<tr>
<td></td>
<td>High risk/poor applicability</td>
<td>0 (0)</td>
</tr>
<tr>
<td></td>
<td>Unclear</td>
<td>0 (0)</td>
</tr>
</tbody>
</table>

CAGI = Canadian Adolescent Gambling Inventory; DSM-IV-J/DSM-IV-MR-J = the Diagnostic Statistical Manual IV (Multiple Response format) adapted for Juveniles; GABSA = Gambling Addictive Behavior Scale for Adolescents; GPSS = Gambling Problem Severity Subscale; MAGS = Massachusetts Gambling Screen; SOGS-RA = the South Oaks Gambling Screen—Revised for Adolescents.

Risk of bias was assessed with the following questions:

* Were participants selected randomly? Was the sample representative of the general population? Were inappropriate exclusions avoided?
* Did the conduct of the test avoid introducing bias?
* Were all participants included in the analysis? Applicability to the review was assessed with the questions:
* Does the sample match the review question?
* Is the validity of the index test a research question of the study?
* Is a reference standard used to make judgments about the validity or reliability of the index test?
Table 3
Characteristics, applicability, and risk of bias of included studies on instruments for measurement of gambling problems among young people

<table>
<thead>
<tr>
<th>Ref. Country (year)</th>
<th>Instrument (classification cutoff score)</th>
<th>Study characteristics</th>
<th>Sample characteristics</th>
<th>Applicability</th>
<th>Risk of bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>[28] USA (2009)</td>
<td>SOGS-RA; DSM-IV-MR-J; DIS-IV adapted for adolescents (≥3 gambling problems in the past year = gambling problems)</td>
<td>a) Cross-sectional</td>
<td>n = 2,274 (4,467) female: 49.5% age: (14–21)</td>
<td>Good</td>
<td>Low</td>
</tr>
<tr>
<td>[30] USA (2009)</td>
<td>MSOGST (1–4 = at risk; ≥5 = probable pathological)</td>
<td>a) Experimental</td>
<td>n = 8,455 female: 52% age: not specified</td>
<td>Unclear</td>
<td>High</td>
</tr>
<tr>
<td>[31] Canada (2009)</td>
<td>SOGS-RA; SOGS (no categories used)</td>
<td>a) Longitudinal</td>
<td>n = sample A: 502 (1,037); sample B: 663 (2,000) female: 0% age: middle childhood 16.2/16.2; early adulthood 22.8/22.5</td>
<td>Good</td>
<td>Low</td>
</tr>
<tr>
<td>[33] USA (2010)</td>
<td>SOGS-RA (≥2 symptoms)</td>
<td>a) Convergent validity with DSM instrument</td>
<td>n = 1,000 (2,274) females: 51.5% age: 18–21</td>
<td>Good</td>
<td>Low</td>
</tr>
<tr>
<td>[34] Norway (2010)</td>
<td>SOGS-RA (2–3 = at risk; ≥4 = problem gambling)</td>
<td>a) Cross-sectional over three time waves (2004; 2005; 2006)</td>
<td>n = 20,648; 21,260; 20,573 (25,037; 24,560; 24,137) female: 50.1%; 49.9%; 50.4% age: 15 (13–19)</td>
<td>Good</td>
<td>Low</td>
</tr>
<tr>
<td>[35] USA (2011)</td>
<td>SOGS-RA; DSM-IV-MR-J; DIS-IV (≥3 gambling problems in the past year = gambling problems)</td>
<td>a) Cross-sectional</td>
<td>n = 2,258 (2,274) female: 49.5% (total sample) age: (14–21)</td>
<td>Good</td>
<td>Low</td>
</tr>
<tr>
<td>[36] Canada (2011)</td>
<td>SOGS-RA; SOGS (no categories used)</td>
<td>a) Longitudinal</td>
<td>n = 1,004 (1,162) female: 0% age: measurements at age 10, 14, 17 (SOGS-RA), 23 (SOGS)</td>
<td>Poor</td>
<td>High</td>
</tr>
<tr>
<td>[37] Italy (2011)</td>
<td>SOGS-RA (≥5 = probable pathologic gambling)</td>
<td>a) Cross-sectional</td>
<td>n = 2,853</td>
<td>Good</td>
<td>High</td>
</tr>
<tr>
<td>[38] USA (2011)</td>
<td>SOGS-RA (≥2 = at risk or problem gambler)</td>
<td>a) Cross-sectional</td>
<td>n = 634 (749) female: 37.1% age: 15.8 (SD 1.4)</td>
<td>Poor</td>
<td>High</td>
</tr>
<tr>
<td>[39] USA (2012)</td>
<td>SOGS-RA (no categories used)</td>
<td>a) Cross-sectional</td>
<td>n = 48 (56) female: 37% age: 21.1 (SD = 2.2; participants aged ≥30 excluded)</td>
<td>Poor</td>
<td>Low</td>
</tr>
<tr>
<td>Ref.</td>
<td>Country (year)</td>
<td>Instrument (classification cutoff score)</td>
<td>Study characteristics</td>
<td>Sample characteristics</td>
<td>Applicability</td>
</tr>
<tr>
<td>------</td>
<td>----------------</td>
<td>------------------------------------------</td>
<td>----------------------</td>
<td>------------------------</td>
<td>--------------</td>
</tr>
</tbody>
</table>
| [40] | USA (2012)     | SOGS-RA (2–3 = at risk; ≥4 = problem gambling/ probable pathological) | a) Cross-sectional  
b) Sample from three selected local high schools  
c) Self-report questionnaire  
d) Cronbach alpha | n = 183 (192)  
female: 48.4%  
age: 15.9 (13–19) | Good  
Poor  
High | Low  
Low  
Low |
| [41] | USA (2012)     | SOGS-RA (2–3 = at risk; ≥4 = problem) | a) Longitudinal  
b) Random sampling from local primary schools  
c) Self-reported questionnaire  
d) Cronbach alpha | n = 310 (678)  
female: 0%  
age: 11–15 (impulsivity) 17, 19, 20 (SOGS-RA; highest score used for analyses) | Poor  
Low  
Unclear | Low  
Low  
Low |
| [42] | Germany (2012) | SOGS-RA (2–3 = at risk; 4–5 = problem; ≥5 = probable pathological gambling; classifications combined for analysis) | a) Cross-sectional  
b) Randomly selected local high school sample  
c) Self-report questionnaire  
d) Cronbach alpha | n = 2,553 (2,640)  
female: 49.3%  
age: 16.7 (12–25) | Good  
Low  
Low | Poor  
Low  
Low |
b) Randomly selected local high school sample  
c) Self-report questionnaire  
d) Cronbach alpha  
e) Correlations between scales | n = 258  
female: 25.2%  
age: 16.13 (12–19) | Good  
Poor  
High | Low  
Low  
Low |
| [44] | Italy (2013)    | SOGS-RA (2–3 = at risk; ≥4 = problem gambler) | a) Cross-sectional  
b) Randomly selected local high school sample  
c) Self-report  
d) Test information function | n = 871 (981)  
females: 36% (total sample)  
age: 16.57 (14–20) | Good  
Low  
Low | Poor  
Low  
Low |
| [45] | Italy (2013)    | SOGS-RA (2–3 = at risk; ≥4 = problem gambler) | a) Cross-sectional  
b) Nationwide high school sample  
c) Self-report  
d) Correlated Multiple Correspondence Analysis, kappa coefficient, Cronbach alphas, item endorsement factor structure | n = 5,930 (n = 14,910)  
females: 48.6% (total sample)  
age: 17 (15–19) | Good  
Low  
Low | High  
Low  
Low |
| [46] | Italy (2013)    | SOGS-RA (broad definition see [47]): no problem, at risk, problem gambling) | a) Cross-sectional  
b) Randomly selected local high school sample  
c) In-class self-report questionnaire  
d) Cronbach alpha  
e) Gambling frequency | n = 943 (994)  
females: 46% (total sample)  
age: 16.57 (first- to fifth-year students) | Good  
Low  
Low | Poor  
Low  
Low |
| [48] | USA (2013)      | SOGS-RA; SOGS (2–3 = at risk; ≥4 = problem; cumulative measures made for gambling variable by using participants' highest involvement at any year.) | a) Longitudinal  
b) Random sampling from local primary schools  
c) Self-reported questionnaire  
d) Cronbach alpha | n = 514 (678)  
female: 47.1% (total sample)  
age: 11–14 (parental monitoring) 16, 18, 19 (SOGS-RA), 20–22 (SOGS) | Poor  
Low  
Unclear | Low  
Low  
Low |
| [49] | Canada (2013)   | DSM-IV-J (2–3 = at risk; ≥4 = probable pathological gambling) | a) Cross-sectional  
b) Community-based high school sample  
c) Self-report questionnaire  
d) Cronbach alpha  
e) Correlation between instruments | n = 532 (2,004)  
female: 36.5%  
age: 16.29 (14–18) | Good  
Low  
Unclear | High  
Low  
Unclear |
| [50] | Italy (2013)    | SOGS-RA (Broad definition see [47]) no problem, at risk, problem gambling); Gambling Attitude Scale | a) Cross-sectional  
b) Sample from four local high schools  
c) Self-report questionnaire  
d) Cronbach alpha | n = 960 (981)  
female: 36% (total sample)  
age: 16.57 (13–23) | Good  
Low  
Low | High  
Low  
Low |
| [51] | Switzerland (2013) | SOGS-RA (French version; adapted: 8 of 12 items used) (2–3 = at risk; ≥4 = problem gambling) | a) Cross-sectional  
b) Local noncompulsory secondary school sample  
c) Online questionnaire  
d) Cronbach alpha | n = 1,102 (1,126)  
female: 48.7%  
age: 15–20 (73.7% under 18 years of age) | Good  
Low  
Low | High  
Low  
Low |
| [52] | USA (2013)      | SOGS-RA (≥2 = problem with gambling) | a) Cross-sectional  
b) Sampled from two high schools in different regions  
c) Online self-report questionnaire  
d) Cronbach alpha | n = 743  
female: 57.9%  
age: 18.7 (18–20) | Good  
Low  
Unclear | High  
Low  
Low |
| [53] | Canada (2014)   | SOGS-RA (six items; ≥2 = gambling problem) | a) Cross-sectional survey over several time waves  
b) Two-stage stratified cluster selection school samples in Ontario (classes selected randomly)  
c) Self-report questionnaires  
d) Cronbach alpha | n = 4,000 to 10,000 (range) age: (12–18)  
2013: n = 10,272 (10,398) females: 48.2%  
age: (grades 7–12) | Good  
Low  
Low | Low  
Low  
Low |

(continued on next page)
Table 3 Continued

<table>
<thead>
<tr>
<th>Ref. Country (year)</th>
<th>Instrument (classification cutoff score)</th>
<th>Study characteristics$^a$</th>
<th>Sample characteristics$^b$</th>
<th>Applicability$^c$</th>
<th>Risk of bias$^d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>[54] USA (2014)</td>
<td>SOGS-RA (no categories used)</td>
<td>a) Longitudinal</td>
<td>n = 515 (678) females: 45% age: 17–22</td>
<td>Good</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b) Randomized block design of local schools</td>
<td></td>
<td>Poor</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c) Self-reported questionnaire</td>
<td></td>
<td>Poor</td>
<td>Low</td>
</tr>
<tr>
<td>[55] Canada (2014)</td>
<td>SOGS-RA (six items; ≥2 = gambling problem)</td>
<td>a) Cross-sectional</td>
<td>n = 4,851 (4,980) females: 53% (total sample) age: 14.6 (grades 7–12)</td>
<td>Good</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b) Stratified cluster sample of Ontario students</td>
<td></td>
<td>Poor</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c) Self-report questionnaires</td>
<td></td>
<td>Poor</td>
<td>Unclear</td>
</tr>
<tr>
<td>[56] Italy (2014)</td>
<td>SOGS-RA (broad definition see [47]): no problem, at risk, problem gambling)</td>
<td>a) Experimental</td>
<td>n = 181 females: 36% age: 15.95 (15–18)</td>
<td>Good</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b) Sample from two local high schools</td>
<td></td>
<td>Poor</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c) In-class self-report questionnaire before and after intervention (McNemar tests)</td>
<td></td>
<td>Poor</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b) Local school sample</td>
<td></td>
<td>Poor</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c) In-class self-report questionnaire</td>
<td></td>
<td>Poor</td>
<td>Low</td>
</tr>
<tr>
<td>[58] Canada (2015)</td>
<td>SOGS-RA; SOGS (1–4 = some problems with gambling; ≥5 = probable pathologic gambler)</td>
<td>a) Longitudinal</td>
<td>At ages 15, 22, and 30: n = 1,882, 1,785, and 1,358 females: 50.2%, 55.3%, and 59.8% age: 15, 22, 30</td>
<td>Good</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b) Random stratified sample of Quebec students</td>
<td></td>
<td>Poor</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c) Structured interview (age 15 and 22), self-report questionnaire (age 30)</td>
<td></td>
<td>Poor</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b) Convenience school sample</td>
<td></td>
<td>Poor</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c) Self-report online questionnaire completed at school computer laboratory</td>
<td></td>
<td>Poor</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td></td>
<td>d) Cronbach alpha</td>
<td></td>
<td>Poor</td>
<td>Low</td>
</tr>
<tr>
<td>[60] Spain (2015)</td>
<td>SOGS-RA (≥2 = at risk; ≥4 = problem)</td>
<td>a) Cross-sectional</td>
<td>n = 1,447 female: 44.9% age: 12.8 (11–16)</td>
<td>Good</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b) Random clustered sample from 17 randomly selected schools</td>
<td></td>
<td>Poor</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c) In-class self-report questionnaire</td>
<td></td>
<td>Poor</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td></td>
<td>d) Cronbach alpha</td>
<td></td>
<td>Poor</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b) Random school sample</td>
<td></td>
<td>Poor</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c) Self-report questionnaire</td>
<td></td>
<td>Poor</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b) Representative sample from six schools</td>
<td></td>
<td>Poor</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c) Self-report questionnaire</td>
<td></td>
<td>Poor</td>
<td>Low</td>
</tr>
<tr>
<td>[63] Lithuania (2009)</td>
<td>DSM-IV-MR-J (2–3 = at risk; ≥4 = problem gambling); SOGS-RA (2–3 = at risk; ≥4 = problem gambling)</td>
<td>a) Cross-sectional</td>
<td>n = 835 female: 52.7% age: 14.5 (10–18)</td>
<td>Good</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b) Sample selected from all schools within city of Kaunas</td>
<td></td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>c) Self-report questionnaire</td>
<td></td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>d) Cronbach alpha, classification accuracy, discrimination function analysis</td>
<td></td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>e) Comparison to external references, correlation between instruments</td>
<td></td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>b) Sample of secondary schools from metropolitan and regional areas of Victoria</td>
<td></td>
<td>Poor</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c) Self-report questionnaire</td>
<td></td>
<td>Poor</td>
<td>Low</td>
</tr>
<tr>
<td>[65] Canada (2012)</td>
<td>DSM-IV-MR-J (2–3 = at-risk; ≥4 = probable pathologic gambling)</td>
<td>a) Cross-sectional</td>
<td>n = 1,870 females: 54.1% (total sample) age: 15.43 (14–18)</td>
<td>Good</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b) Convenience sample from local schools</td>
<td></td>
<td>Poor</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c) Self-report questionnaire</td>
<td></td>
<td>Poor</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b) Entire student population aged 12–19 on island of Kos</td>
<td></td>
<td>Poor</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c) Self-report questionnaire</td>
<td></td>
<td>Poor</td>
<td>Low</td>
</tr>
<tr>
<td>[67] England (2013)</td>
<td>NL-CLIP (0–2 = nonproblem); DSM-IV-MR-J (2–3 = at-risk; ≥4 = problem gambling)</td>
<td>a) Cross-sectional</td>
<td>n = 1,425 (8,958) females: 49.6% (total sample) age: 11–15</td>
<td>Good</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b) Schools sampled from England, Scotland, and Wales, classes selected randomly</td>
<td></td>
<td>Poor</td>
<td>Unclear</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c) Self-report questionnaire</td>
<td></td>
<td>Poor</td>
<td>Unclear</td>
</tr>
<tr>
<td></td>
<td></td>
<td>d) Cohen kappa</td>
<td></td>
<td>Poor</td>
<td>Unclear</td>
</tr>
<tr>
<td>Ref. Country (year)</td>
<td>Instrument (classification cutoff score)</td>
<td>Study characteristics</td>
<td>Sample characteristics</td>
<td>Applicability</td>
<td>Risk of bias</td>
</tr>
<tr>
<td>--------------------</td>
<td>------------------------------------------</td>
<td>----------------------</td>
<td>-----------------------</td>
<td>--------------</td>
<td>-------------</td>
</tr>
</tbody>
</table>
| Australia (2013)   | Victorian Gambling Screen (0–7 = nonproblem gambling; 8–20 = borderline problem; ≥21 = problem gambling); DSM-IV-J (≥4 = pathologic gambling) | a) Cross-sectional  
   b) Sample from 18 schools selected from Australian Capital Territory  
   c) Self-report questionnaire (not directly specified)  
   d) Cronbach alpha  
   e) Instrument correlation, classification comparison between instruments | n = 926 female: 48.4% age: 14.46 (approximately 11–19) | Good  
   Poor  
   Good  
   Good  
   Unclear  |
| China (2014)       | DSM-IV-J (2–3 = at risk; ≥4 = probable pathologic gambling) | a) Cross-sectional  
   b) Random sampling of high schools in Hong Kong  
   c) Self-report questionnaire | n = 4,734 (5,523) females: 49.3% age: 16.39 (12–23) | Good  
   Poor  
   Poor  
   Poor  |
| Canada (2014)      | DSM-IV-J (2–3 = at risk, ≥4 = problem gambling); GRCS | a) Cross-sectional  
   b) Several high schools sampled from a school district  
   c) Self-report questionnaire  
   d) Cronbach alpha, DSM-IV-J classification predicted with GRCS scores | n = 1,490 female: 57.7% age: 17.10 (16–18) | Good  
   Poor  
   Poor  
   Poor  |
| Canada (2014)      | DSM-IV-J (2–3 = at risk; ≥4 = problem gambling); GRCS | a) Cross-sectional  
   b) Several secondary schools sampled from a school district  
   c) Self-report questionnaire  
   d) Cronbach alpha  
   e) Cronbach alpha, DSM-IV-J classification predicted with GRCS scores | n = 2,004 female: 57.7% age: 16.51 (14–18) | Good  
   Poor  
   Poor  
   Poor  |
| Finland (2015)     | DSM-IV-MR-J (≥2 = at risk and problem gambling) | a) Cross-sectional  
   b) Convenience sample from 11 schools  
   c) In-class self-report questionnaire  
   d) Cronbach alpha, factor analysis, classification accuracy  
   e) Gambling frequency | n = 988 female: 46.8% age: 13.41 (12–15) | Good  
   Good  
   Good  
   Good  |
| Israel (2015)      | DSM-IV-MR-J (2–3 = at risk; ≥4 = probable pathologic gambling) | a) Cross-sectional  
   b) Convenience sample from six schools  
   c) In-class self-report questionnaire  
   d) Cronbach alpha | n = 595 female: 60% age: 15.13 (13–19) | Good  
   Poor  
   Poor  |
| Norway (2009)      | MAGS (3–4.5 = problem gambler, ≥5 = pathologic gambler) | a) Cross-sectional  
   b) Randomly selected (at the class level) high school sample  
   c) Online self-report questionnaire during designated school time  
   d) Item response theory: Differential Item Functioning | n = 1,285 (1,351) female: 47.5% age: 17.3 (16–19) | Good  
   Good  
   Poor  |
| USA (2014)         | MAGS (used according to DSM-IV and DSM-5 criteria) | a) Cross-sectional  
   b) Convenience sample from 10 local high schools  
   c) Self-report questionnaire  
   d) Latent class analysis  
   e) OR for spending >1 hour/week gambling | n = 3,901 (4,523) female: 51.5% age: (<14–18) | Good  
   Good  
   Good  |
| USA (2015)         | MAGS (≥1 = at risk and problem gambling) | a) Cross-sectional  
   b) Convenience high school sample  
   c) Administration method not specified  
   d) Cronbach alpha | n = 1,988 female: 39.2% age: (9th to 12th grade) | Good  
   Poor  
   Unclear  
   High  |
| Canada (2010)      | CAGI (GPSS 0–1 = no problem; 2–5 = low-to-moderate severity; ≥6 = high severity); DSM-IV (1–3 = low gambling problem; ≥4 = problem gambling); SOGS-RA (2–3 = at-risk gambling; ≥4 = problem gambling); CRAGS [four classes; ≥5 = problem gambling] | a) Cross-sectional  
   b) Phase II sample from schools, Phase III sample from local substance abuse and detention treatment centers  
   c) Self-report questionnaires; clinical interviews  
   d) Cronbach alpha, Intraclass correlations, principal component analysis, confirmatory factor analysis, discriminant function analysis, ROC analysis  
   e) Correlations with references | Phase II school sample n = 864 | Good  
   Low  
   Good  
   Good  |

(continued on next page)
Table 3 Continued

<table>
<thead>
<tr>
<th>Ref.</th>
<th>Country (year)</th>
<th>Instrument (classification cutoff score)</th>
<th>Study characteristicsa</th>
<th>Sample characteristicsb</th>
<th>Applicabilityc</th>
<th>Risk of biasd</th>
</tr>
</thead>
<tbody>
<tr>
<td>[79]</td>
<td>South Korea (2012)</td>
<td>Gambling Addictive Behavior Scale for Adolescents (no classifications specified)</td>
<td>a) Cross-sectional b) Convenience stratified school sample classes selected randomly c) Self-report questionnaire d) Cronbach alpha, principal component analysis e) Correlation with other variables</td>
<td>n = 299 (320) female: 40.8% age: not specified</td>
<td>Good</td>
<td>Good</td>
</tr>
</tbody>
</table>

CAGI = Canadian Adolescent Gambling Inventory; CRAGS = Clinician Rating of Adolescent’s Gambling Severity; DIS-IV = Diagnostic Interview Schedule, Version IV; DSM-IV-J (DSM-IV-(MR)-J) = the Diagnostic Statistical Manual IV (Multiple Response format) adapted for Juveniles; GRCS = Gambling Problem Severity Subscale; GRCS = Gambling-Related Cognitions Scale; MAGS = Massachusetts Gambling Screen; MSOGST = Modified South Oaks Gambling Screen for Teens; NL-CLIP = short screen for problem gambling among children based on criteria identified in NODS-CLIP (NODS-CLIP: Diagnostic Screening for Gambling Disorders [NODS] Loss of Control, Lying and Preoccupation [CLIP]); ROC = receiver operating characteristic; SOGS-RA = the South Oaks Gambling Screen—Revised for Adolescents.

a) Design, b) sampling method, c) administration method, d) statistical methods related to reliability and validity, e) references of ARPG; comparator measurement.

b) Sample size in analysis and in total, gender distribution from analyzed sample (unless otherwise specified as total sample), mean age (range) in years.

c) QUADAS-2 assessment for applicability concerns: patient selection, index test, reference standard.

d) QUADAS-2 assessment of risk of bias concerning patient selection, index test, and flow and timing.

Finland, Germany, Greece, Hong Kong, Israel, Lithuania, Malaysia, South Korea, Spain, and Switzerland.

Table 2 summarizes the risk of bias in the articles. Almost half (n = 24) of the articles had done patient selection well. Most (n = 38) had conducted the index test properly, and in nearly all articles (n = 43), participants were maintained in analyses.

The applicability of the articles to this review was more variable (Table 2). In most (n = 43) of the articles, patient selection suited our study question, whereas only some had good applicability related to evaluating the index test (n = 9) or for using a reference measurement (n = 12).

South Oaks Gambling Screen Revised for Adolescents

In articles evaluating the SOGS-RA as the primary adolescent ARPG instrument, high risk of bias was often evident for patient selection but not for other items. Although the sample was often appropriate, only a few of the SOGS-RA articles primarily investigated instrument properties or used a reference test (Table 2).

Estimates of internal consistency (Cronbach’s $\alpha$) for the SOGS-RA ranged from .60 [54] to .94 [49] ($M = .76$). The instrument seems stable over time [45]. Cronbach $\alpha$ for the Modified South Oaks Gambling Screen for Teens (MSOGST) was .87 [31]. The MSOGST includes 20 items (12 items used for scoring; score range 0–20) with rewording for adolescents [81].

The SOGS-RA displayed good classification accuracy, with a sensitivity of .97, specificity of .99, false-positive rate of .20, and false-negative rate of .00015 when using the DSM-IV-MR-J as a reference [64]. “Feeling bad about money lost” and “Gambling more than planned” were the most commonly reported items in two studies [44,45], whereas another study found highest rates for items “Gambling more than planned” and “Experiencing problems with family or school due to gambling” [53]. For both items, females were less likely to endorse them than males [53], contradicting the finding that item endorsement differences would not exist between genders [45]. The least often reported items were “Absent from school due to betting” and “Borrowed or stolen for bets or debts” [45]. Items “Lying about winning,” “Gambling more than planned,” “Felt bad about money lost,” and “Hidden any signs of gambling” were more likely endorsed by Asian-Americans than by white Americans [59]. Overall, the SOGS-RA items had good discrimination, and the item severity was appropriate for screening ARPG [44].

Two Italian studies suggest a single factor structure of the SOGS-RA [44,45]. The SOGS-RA correlated positively with the scope of gambling activities ($r = .57$) [29] and gambling frequency ($r = .59$) [38,45]. The SOGS-RA score correlation with gambling related crimes was .26 [25]. Correlations with the SOGS-RA and the DSM-IV-J ($r = .64$) [49] and the DSM-IV-MR-J ($r = .74$–.892) [32,64] indicate convergence. The SOGS-RA correlated positively with the Chinese version of the Gambling Belief Questionnaire [39] and the Gambling Urge Scale ($r = .60$) [39]. In a longitudinal study [31], the SOGS-RA correlated positively with the SOGS; the two were considered metrically invariant, thus psychologically comparable.

Finally, the SOGS-RA indicated sensitivity to change, as the prevalence of ARPG decreased after the removal of banknote acceptors in slot machines, compared to the two preceding baseline years during which ARPG prevalence remained stable [34]. The SOGS-RA also detected a decrease in ARPG prevalence following intervention [56].

DSM-IV-(MR)-J

Of the 12 articles evaluating the DSM-IV-(MR)-J, six had a high risk of bias concerning patient selection, while four instances of unclear risk of bias were evident. Otherwise, the risk of bias was low (Table 2). In patient selection, these studies had good applicability. Two studies investigated instrument properties, and in four cases, a reference standard was used.

Cronbach $\alpha$ ranged from .75 [62] to .93 [45] ($M = .85$). Items on illegal acts, tolerance, loss of control, and lying had the highest endorsement rates and were the most sensitive in identifying ARPG, whereas items on escape, withdrawal, and risking job, education, or relationship were the least sensitive [73]. One study suggested a single-factor structure for the DSM-IV-MR-J [73]. The DSM-IV-MR-J distinguished between at-risk and pathologic gamblers, and the escape item was the best discriminator [64]. The DSM-IV-J correlated significantly with the Victorian Gambling Screen ($r = .65$) [69], classifying 41 participants as problem gamblers (19 unidentified by the Victorian Gambling Screen) [69]. The Gambling-Related Cognitions Scale explained 32% of the DSM-IV-J variance [71]. The DSM-IV-MR-J score was positively associated
Findings of studies on instruments for measurement of gambling problems among young people

<table>
<thead>
<tr>
<th>Ref.</th>
<th>Country (year)</th>
<th>Statistical results and author report of reliability/validity findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>[28]</td>
<td>USA (2009)</td>
<td>Total number of items endorsed (on SOGS-RA, DSM-IV-MR-J, and DIS-IV) analysis ( r = .97 )</td>
</tr>
<tr>
<td>[29]</td>
<td>USA (2009)</td>
<td>SOGS-RA ( \alpha = .85 ); correlation with gambling activities ( r = .57 ) and crime ( r = .26 ) (( p &lt; .001 ))</td>
</tr>
<tr>
<td>[30]</td>
<td>USA (2009)</td>
<td>MSOGST ( \alpha = .87 )</td>
</tr>
<tr>
<td>[31]</td>
<td>Canada (2009)</td>
<td>SOGS-RA ( \alpha = .78 ) (sample A) and ( \alpha = .78 ) (sample B); correlation with SOGS-RA and SOGS ( r = .22 ) (( p &lt; .05 )) and ( .28 ) (( p &lt; .05 )); SOGS and SOGS-RA were metrically invariant, and thus psychologically comparable.</td>
</tr>
<tr>
<td>[32]</td>
<td>USA (2009)</td>
<td>SOGS-RA ( \alpha = .74 ); SOGS-RA correlation with DSM-IV-MR-J ( r = .76 )</td>
</tr>
<tr>
<td>[33]</td>
<td>USA (2010)</td>
<td>SOGS-RA ( \alpha = .74 ) (( n = 2,274 ); age 14–21)</td>
</tr>
<tr>
<td>[34]</td>
<td>Norway (2010)</td>
<td>At-risk and problem gambling displayed stability during 2004 and 2005 (preintervention), with a significant decrease after the removal bank note acceptors in 2006 (postintervention). Thus, the SOGS-RA displayed stability and sensitivity to change. No gender- or age-related differences were evident in the reduction of gambling problem prevalence (i.e., 2005–2006).</td>
</tr>
<tr>
<td>[35]</td>
<td>USA (2011)</td>
<td>SOGS-RA ( \alpha = .72 ); DSM-IV-MR-J ( \alpha = .71 ); DIS-IV ( \alpha = .77 ); combined ( \alpha = .89 )</td>
</tr>
<tr>
<td>[36]</td>
<td>Canada (2011)</td>
<td>SOGS-RA ( \alpha = .76 )</td>
</tr>
<tr>
<td>[37]</td>
<td>Italy (2011)</td>
<td>SOGS-RA ( \alpha = .80 )</td>
</tr>
<tr>
<td>[38]</td>
<td>USA (2011)</td>
<td>SOGS-RA ( \alpha = .83 ); SOGS-RA correlation with gambling frequency ( r = .59 )</td>
</tr>
<tr>
<td>[39]</td>
<td>USA (2012)</td>
<td>SOGS-RA ( \alpha = .84 ); Gambling Urge Scale and SOGS-RA correlation ( r = .60 ) (( p &lt; .001 )); Postexposure Gambling Urge Scale scores correlated significantly with SOGS-RA scores</td>
</tr>
<tr>
<td>[40]</td>
<td>USA (2012)</td>
<td>SOGS-RA ( \alpha = .80 )</td>
</tr>
<tr>
<td>[41]</td>
<td>USA (2012)</td>
<td>SOGS-RA ( \alpha = .71 )</td>
</tr>
<tr>
<td>[42]</td>
<td>Germany (2012)</td>
<td>SOGS-RA ( \alpha = .77 )</td>
</tr>
<tr>
<td>[43]</td>
<td>Hong Kong (2012)</td>
<td>Chinese version of the Gamblers Belief Questionnaire ( \alpha = .91 ); Chinese version of the Gamblers Belief Questionnaire correlation range with other scales ( .40–.75 )</td>
</tr>
<tr>
<td>[44]</td>
<td>Italy (2013)</td>
<td>Factor loadings range (.53–.83) (( p &lt; .001 )); CFI = .96; TLI = .97; RMSEA = .03; Suggests single factor structure for SOGS-RA. Items “feeling bad about money lost” and “gambling more than planned” had the highest endorsement rate. Absence from school due to betting, borrowing money, and stealing for betting were the most discriminative. Items “lying about winning” and “wanting to stop gambling” had the lowest discrimination. Majority of items had good discrimination. Screen accurately measures medium to high levels of problem gambling (i.e., item severity located along intended range).</td>
</tr>
<tr>
<td>[45]</td>
<td>Italy (2013)</td>
<td>Multiple correspondence analysis principle component (eigenvalue = 3.875) explained 32.3% of variance; test–retest ( r = .70 ); suggested single factor structure for SOGS-RA. The least endorsed items were “absent from school due to betting” and “borrowed or stolen for bets or debts.” The most endorsed items were “gambling more than intended” and “feeling bad about the amount bet.” The SOGS-RA score was positively associated to gambling frequency. No gender differences were evident in item endorsements. The SOGS-RA seems to be stable over time.</td>
</tr>
<tr>
<td>[46]</td>
<td>Italy (2013)</td>
<td>SOGS-RA ( \alpha = .73 )</td>
</tr>
<tr>
<td>[47]</td>
<td>USA (2013)</td>
<td>SOGS-RA ( \alpha = .61–.72 )</td>
</tr>
<tr>
<td>[48]</td>
<td>Canada (2013)</td>
<td>SOGS-RA ( \alpha = .94 ); DSM-IV-J ( \alpha = .93 ); SOGS-RA and DSM-IV-J correlation ( r = .64 ); DSM-IV-J is a more conservative instrument for measuring pathologic gambling than SOGS-RA.</td>
</tr>
<tr>
<td>[49]</td>
<td>Italy (2013)</td>
<td>SOGS-RA ( \alpha = .73 ); Gambling Attitude Scale ( \alpha = .80 ); Gambling Attitude Scale discriminated nonproblem gamblers (more cautious perception) from at-risk and problem gamblers. Problem gamblers scored higher on the items from the profitability factor than at-risk and nonproblem gamblers.</td>
</tr>
<tr>
<td>[50]</td>
<td>Switzerland (2013)</td>
<td>SOGS-RA (eight items) ( \alpha = .70 )</td>
</tr>
<tr>
<td>[51]</td>
<td>USA (2013)</td>
<td>SOGS-RA ( \alpha = .82 )</td>
</tr>
<tr>
<td>[52]</td>
<td>Canada (2014)</td>
<td>SOGS-RA (6 items) ( \alpha = .70 ); Most endorsed items in 2013 were “gambling more than planned” (1.7%) and “experiencing problems with family or school due to gambling” (1.5%). Females were significantly less likely than males to endorse the previously mentioned items (( p &lt; .05 )).</td>
</tr>
<tr>
<td>[53]</td>
<td>USA (2014)</td>
<td>SOGS-RA ( \alpha ) range during different years of administration (.60–.72 )</td>
</tr>
<tr>
<td>[54]</td>
<td>Canada (2014)</td>
<td>SOGS-RA (six items) ( \alpha = .71 ); AUC = .80 (concordance of short version and full version of SOGS-RA); The short version of the SOGS-RA may overestimate prevalence rates.</td>
</tr>
<tr>
<td>[55]</td>
<td>Italy (2014)</td>
<td>McFadden’s ( \chi^2 ) (1, ( N = 88 )) = 8.77, ( p &lt; .05 ); At baseline the prevalence of ARPG (measured with SOGS-RA) in the training group was 41%. After training (approximately 6 months after pretest) a significant and medium in size reduction in the prevalence of ARPG was evident (prevalence at follow up: 28%). The intervention intended to increase correct knowledge about gambling, reduce misconceptions, economic optimistic view of gambling profitability, and superstitious beliefs. The intervention contained activities, PowerPoint slides, a video, and collective discussions.</td>
</tr>
<tr>
<td>[56]</td>
<td>Italy (2015)</td>
<td>SOGS-RA ( \alpha = .73 ) (CI = .70–.75); SOGS-RA ( \alpha = .76 ); omega = .68; Asian-Americans were more likely to endorse the following items compared to white Americans: lying about winning; gambling more than intended; feeling bad about money bet; hidden signs of gambling.</td>
</tr>
<tr>
<td>[57]</td>
<td>Spain (2015)</td>
<td>SOGS-RA ( \alpha = .83 )</td>
</tr>
<tr>
<td>[58]</td>
<td>Malaysia (2015)</td>
<td>SOGS-RA ( \alpha = .77 )</td>
</tr>
<tr>
<td>[59]</td>
<td>Australia (2009)</td>
<td>DSM-IV-J ( \alpha = .82 )</td>
</tr>
<tr>
<td>[60]</td>
<td>Lithuania (2009)</td>
<td>DSM ( \alpha = .80 ); SOGS-RA ( \alpha = .75 ); DSM-IV-MR-J and SOGS-RA correlation ( r = .892 ) (( p &lt; .001 )); SOGS-RA ( \kappa = .833 ) (( p &lt; .001 )); SOGS-RA is more liberal in classifying gambling pathology. SOGS-RA classified 34 of 35 pathologic gamblers correctly, using the DSM-IV-MR-J as the reference. DSM distinguished between social, at-risk, and pathologic gamblers. DSM item on “Escape” identified as best discriminator. Those who gambled at least 1 per week scored significantly higher than participants who gambled more seldom. SOGS-RA sensitivity = .97 (34/35), specificity .986, false-positive rate = .20 and false-negative rate = .00015. DSM pathologic gamblers were likely to spend more money on gambling than nonpathologic gamblers.</td>
</tr>
<tr>
<td>[61]</td>
<td>Australia (2010)</td>
<td>DSM-IV-MR-J ( \alpha = .78 )</td>
</tr>
<tr>
<td>[62]</td>
<td>Canada (2012)</td>
<td>DSM-IV-MR-J ( \alpha = .75 )</td>
</tr>
</tbody>
</table>

(continued on next page)
with gambling frequency [73]. In line with previous research, the DSM-IV-J was more conservative than the SOGS-RA [49].

**Massachusetts Gambling Screen**

Three articles concerning the MAGS had low risk of bias for patient selection, whereas two articles had low risk of bias for both the conduct of the index test and flow and timing. Two of the articles had high applicability for the index test (Table 2).

Cronbach $\alpha$ for the MAGS was .92 [77]. Two studies [75,76] were concerned with instrument psychometric properties. ARPG classification was associated to a stronger likelihood of gambling at least 1–2 hours/week compared to low-risk gamblers [76]. All items had different functioning between genders,
such that males were more likely to endorse each item, given otherwise equal scores for the latent variable [75]. Latent class analysis suggested a four-class solution for the DSM criteria (measured with the MAGS): low-risk gambling, at-risk chasing gambling, at-risk negative consequences gambling, and problem gambling [76].

Canadian Adolescent Gambling Inventory

The report on the CAGI [78] had low risk of bias on all domains and good overall applicability. It is a comprehensive assessment of the scale beyond the scope of this systematic review, so we provide a brief summary. Cronbach’s α for the subscales ranged from .83 to .90. The correlation with the DSM-IV criteria and the Clinician Rating of Adolescent’s Gambling Severity was ≥.89, and correlation with gambling involvement measures ranged from .14 to .67. Principal component analysis and confirmatory factor analysis suggested a four-factor solution (psychological consequences, social consequences, financial consequences, and loss of control). The nine-item Gambling Problem Severity Subscale (GPSS) was optimal for classification performance as defined by discriminant function analysis and receiver operating characteristic (ROC) analysis.

Gambling Addictive Behavior Scale for Adolescents

The article on the GABSA [79] had an unclear risk of bias for conducting the index test but otherwise low risk and overall good applicability.

The developmental report of the GABSA suggested the scale to have high internal consistency, with an overall Cronbach’s α of 94, and subscale α ranging from .88 to .90. The scale displayed significant (p ≤ .001) classification accuracy (area under ROC curve analysis [AUC]) of .71 for the nongambling group, .75 for the nonproblem gamblers, and .74 for the problem gamblers, using reference scores from the Addictive Personality subscale of the Eysenck Personality Questionnaire. As for the psychometric properties, the scale loaded onto four factors (loss of control, life dysfunction from gambling addiction, gambling experience, and social dysfunction from problem gambling) explaining 55% of variance. The scale further displayed convergent validity by significant positive correlations with addictive personality, irrational gambling beliefs, and gambling behavior.

Discussion

Our review aimed to clarify which instruments measuring ARPG among youth are reliable and valid in light of reported estimates of internal consistency, classification accuracy, and psychometric properties. Five ARPG instruments were examined: 3 of the 33 SOGS-RA articles and 2 of the 12 DSM-IV-(MR)-J articles investigated instrument properties. Two of the three MAGS articles looked at instrument properties. The remaining articles concerned the development of the CAGI and the GABSA.

South Oaks Gambling Screen Revised for Adolescents

Most (58%) of the SOGS-RA articles had potentially biased sampling procedures. The reviewed estimates for internal consistency were similar to earlier findings [10,13,82], although variation was evident. Estimates of correlations with the SOGS-RA and external references (e.g., gambling frequency) and the corresponding DSM-IV-(MR)-J instruments parallel previous findings [5,10].

Specific items are problematic in population-based studies because of uneven endorsement rates. Controversially, there was evidence suggesting no item endorsement differences between genders [45] but also that females are less likely to endorse items than males [53]. The SOGS-RA was psychologically parallel to the SOGS [31], contradicting the idea of uniqueness of youth gambling compared to adult gambling. The results on the SOGS-RA highlight the lack of a unique conceptualization of gambling problems for adolescents and are inconclusive concerning gender differences [82].

Almost half of the SOGS-RA articles used ethnically diverse samples, consisting African-Americans, Hispanics, Asian-Americans, and samples from Hong Kong and Malaysia. Although these studies were not primarily testing the SOGS-RA, studies with different ethnicities emerged, as previously deemed necessary [1]. Cronbach’s α estimates among samples with high rates of African-Americans ranged from .60 to .85, which is moderate at best, but within the same range as previous estimates. Two items endorsed more often by Asian-Americans than white Americans concerned minimizing other people’s worries about the respondents gambling, suggesting a difference in attitudes toward the social aspect of gambling [59].

DSM-IV-(MR)-J

Estimates of internal consistency parallel previous findings of about .80 [15,16], which is considered satisfactory. Two DSM-IV-MR-J articles were primarily concerned with instrument properties and indicated that items are not equivalent in their ability to detect problem gambling, and variabilities in item endorsement rates and sensitivity are evident [64,73]. Gambling frequency, amount of money spent on gambling, and score on other problem gambling screening instruments were positively associated to the DSM-IV-MR-J score [64,69,73], in alignment with previous findings [5,83]. Agreement between the SOGS-RA and the DSM-IV-MR-J has previously been found to be greater among males than females [5]. Corresponding items of the DSM-IV-MR-J and the NLCliP (short screen for problem gambling among children based on the criteria identified in the NODS-CLiP [Diagnostic Screening for Gambling Disorders - Loss of Control, Lying, and Preoccupation]) had deviating endorsement rates [69], highlighting the importance of item wording, as misinterpretation may magnify prevalence rates [14,68]. The concept of gambling to escape a negative state of mind may be foreign to and poorly understood by young individuals, leading to false-positive responses [84].

Massachusetts Gambling Screen

This review provided one estimate of internal consistency (α = .92) [77] but no indices for classification accuracy of the MAGS DSM-IV subscale and no information on the MAGS subscale. The seven-item subscale of the MAGS (MAGS-7) Cronbach α was .83, and the DSM-IV subscale .87 [17]. MAGS-7 correctly classified 96% of adolescents as problem gamblers, at-risk gamblers, and nonproblem gamblers when using the DSM-IV criteria as a reference. The MAGS-7 item “Are you always able to stop gambling when you want” was the least discriminating item, and low congruence for the classification of ARPGers was evident between the MAGS-7 and SOGS-RA [12].
We found new information concerning the psychometric properties of the DSM-IV subscale of the MAGS, suggesting there may be different types of at-risk gamblers: some mostly chase losses, whereas others more notably experience negative consequences. This finding supports the notion of multiple trajectories leading to problem gambling, as suggested by the Pathways Model [85]. The item on illegal acts was not useful for classifying gamblers [76]. Item endorsement differed significantly between genders, suggesting that the validity of the criteria is not equivalent between genders [75].

**Canadian Adolescent Gambling Inventory**

The CAGI was included to this study even though it is not a population-based instrument. During its development, the CAGI was tested with student samples and a clinical sample, all resulting in the same four-factor structure [86]. Cronbach α for the factors ranged from .83 to .90 or within the range often estimated for other instruments. The CAGI displays high congruence to self-rated DSM-IV criteria, clinician-rated DSM-IV criteria, and to the Clinician Rating of Adolescent’s Gambling Severity, although the validity of these comparators is unknown [78].

The administration time of the CAGI (20 minutes) [9] may inhibit its use, but the nine-item GPSS subscale is shorter and asks about frequencies of concrete gambling behaviors. A unique strength of the GPSS subscale is that items inquire about the effects of gambling on peer relationships and financial consequences in a developmentally appropriate manner [84]. However, conclusions on CAGI need to be done cautiously since data are limited to one clinical study which was not reported in a peer-reviewed journal.

**Gambling Addictive Behavior Scale for Adolescents**

The article [79] did not provide directions for scoring or classification thresholds, and the instrument lacks strong theoretical framework. The content of proposed factors is questionable, as similar items are distributed across various factors. For example, the factor named loss of control includes the item “Spend money on gambling without paying for necessary things,” whereas the social dysfunction factor includes the item “Bet money or prizes for gambling beyond my pocket money.” Likewise, items “Spend more and more time gambling,” “Do not leave a place all day to gamble,” “Have little time to play other things or do activities except for gambling,” and “It is hard to pass by a PC room, billiards room, or amusement arcade without stopping by” all load onto different factors. However, items do tap into concrete behaviors.

**Theoretical base of the instruments**

A theoretical framework should be a cornerstone for the instruments. In gambling research, surveys largely use adapted adult instruments, where DSM-criteria represent the gold standard. The DSM criteria were originally formulated for diagnostic purposes, not for classifying individuals in a survey [87]. Although many instruments are based on the DSM-IV criteria [88], they lack sufficient empirical evidence from outside the clinical context [89]. The use of the DSM-IV criteria as the gold standard has been criticized among adults [90]; however, the psychiatric criteria for pathologic gambling have never been clinically tested among adolescents [2]. Moreover, research does not recognize the unique developmental characteristics of young people. Developing the CAGI is a step in the right direction.

Potenza et al. [6] classified participants endorsing one or more criteria as ARPGers [6]. Herein, the criteria for ARPG varied. Two or more criteria were used to define ARPGers, with the exception of few studies using one or more criteria [30,58,63,77], whereas some studies defined only problem gambling [28,35,62,75] and/or pathologic gambling [37,67]. Notably, there are clear differences of at-risk and problem gambling among adolescents, at-risk referring to individuals who are starting to develop a number of gambling-related problems, but do not meet the established criteria for the more severe form of gambling (i.e., gambling disorder) [90].

Current instruments are focused on the negative consequences of gambling, paying less attention to actual behavior [84]. Loss of control, however, is a behavior associated with problem gambling [91] taken into account to some degree in measurement. For example, pursuing lost money indicates loss of control and instruments included items for “chasing losses” [84,92] — this may even be related to a subtype of at-risk gambling occurring in absence of negative consequences of gambling [76]. Another concept that is a fundamental part of addiction, and intuitively linked to loss of control, is craving [93,94]. Craving is not examined in the ARPG instruments.

The conceptualization of problem gambling incorporates biological, psychological, and social aspects [85]. It is a major drawback that theory is not better integrated with measurement of ARPG. The individual-centered perspective to addiction research accounts for individual variance in a range of risk factors (i.e., impulsivity [personality trait] and social factors) [95]. This approach may be helpful to broaden perspectives in gambling research, as the development of gambling problems is influenced by gambling related beliefs, personality traits, and the motivation to gamble [71,96–98].

Different aspects of gambling are not equivalent indicators of problem gambling severity, so imputation of item-weighting procedures is necessary. The adult problem gambling instrument Problem and Pathological Gambling Measure [99] enables the detection of gambling problems even with lack of insight or denial of gambling problems, by formulating items accordingly. These aspects should also be noticed in the measurement of youth ARPG.

**Study limitations**

Most articles had poor applicability concerning the index test and reference measurements; only few had developing ARPG instruments as their primary objective. The low quality of the reviewed literature is a result in itself, but also a limitation. Measuring internal consistency exclusively with Cronbach α may be inappropriate because the assumptions are unlikely to be met [100]. We propose that Tarkkonen rho may be a more appropriate estimate of internal consistency as it applies the general form for estimating reliability, whereas Cronbach α is a special case with restrictive assumptions [100]. Regarding bias in the articles, the overall risk was mainly low. Further studies would benefit from even more in-depth quality assessment including criteria for psychometric properties of the instruments [101,102].

Clinical assessment was used as a comparator only rarely. Gamblers with gambling-related problems have shown more discrepancies between self-report and actual outcomes
(i.e., accuracy bias, underestimation of losses or overestimation of gains) than gamblers without gambling problems [103]. Attention, memory, and other cognitive systems [103], such as already recognized cognitive biases [104,105], might be involved. Thus, in a clinical assessment discussion of possible discrepancies is recommended, as it enables a more detailed evaluation [103]. However, determining the performance of a tool compared to clinical assessment is challenging without a stronger base and defined criteria for ARPG. Without a “gold standard” test, the validation is problematic.

Nearly all articles were population based, so results cannot be generalized into the clinical context [106,107]. Short sensitive measures based on DSM criteria are usually used to screen at-risk individuals in the clinical setting as part of a wider clinical assessment [106]. Screening tools and diagnostic instruments may not be interchangeable as the purposes of the two are fundamentally different [108]; in reality, instruments for assessing gambling are often used outside the context originally intended.

Our material included methodologically diverse studies based on, for example, respondent age, ethnicity, and method of survey administration. To gain a comprehensive understanding of adolescent ARPG, gambling needs to be studied in an ecologically valid manner. This may be achieved with ambulatory methods (i.e., smart phones) [109].

Summary and Implications

An estimate of reliability was reported for five ARPG instruments. Most articles (66%) evaluated SOGS-RA. The GABSA was the only novel instrument developed from 2009. Generally, the evaluation of reliability and validity of the instruments was superficial. Despite a very modest publication base, the CAGI seems to have a strong theoretical and methodological base. Reviewed articles with high applicability [31,44,45,64,73,75,76,78,79] advocate that screening should include measures of risk taking, self-control or impulsivity, delinquent behavior, and social risk factors. The GABSA and the CAGI were the only instruments originally developed especially for youth. Studies were entirely population based, except the one concerning the CAGI. In 2010, Volberg et al. concluded that despite the questions raised regarding the validity of the SOGS-RA and DSM-IV-MR-J, these instruments are the best tools for evaluating adolescent gambling problems while waiting for a better-validated instrument [2]. In 2015, this conclusion still seems up to date.

In the past 5 years, a variety of ethnicities were included in research [1], although cultural validation has not been an actual aim. It is encouraged that future studies look closer into gender differences, study clinical samples, and community-based samples. Attending to these issues in ARPG research may lead to a deeper understanding of the phenomena. In recognition of the similarities across addictions, it is important that accumulating knowledge from the field is integrated into measuring ARPG.

Rigorous psychometric research of youth ARPG instruments, as recommended by earlier reviews [1,9], has not yet been accomplished. Thus, it is untimely to name the most suitable instrument presently available. Researchers are encouraged to test reliability in population-based studies, where samples are not derived from schools, and especially in the clinical context. Reporting alphas from previous articles is not enough, considering the weaknesses in the theoretical foundation of ARPG instruments. We hope that bringing recent work to light will nourish forthcoming studies. If researchers collectively strive to improve the accuracy of ARPG measurement, even the understanding of behavioral addictions in general may advance tremendously.

Acknowledgments

The authors thank the Information Service Designer Pirjo Vuorio, National Institute for Health and Welfare, for her substantial assistance in the management of articles used during the conduct of this systematic review. All authors contributed significantly to the planning of the work and drafting the manuscript or revising the work critically for important intellectual content. The authors of this manuscript and Pirjo Vuorio did not have any interests that might be interpreted as influencing the research. APA ethical standards were followed in the conduct of the study. The corresponding author affirms that all individuals who have contributed significantly to the preparation of this manuscript have been mentioned in the Acknowledgements. Please note that the coauthor Sari Castrén has presented the protocol of our report at the sixth International Gambling Conference in February 2016 (10.-12.2.2016), in Auckland, New Zealand.

Funding Sources

This study was financially supported by the Ministry of Social Affairs and Health, Helsinki, Finland (the 52 Appropriation of the Lotteries Act, contract STM/3189/2011).

Supplementary Data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jadohealth.2016.03.007.

References


R. Edgren et al. / Journal of Adolescent Health 58 (2016) 600–615 613

