Addictive Behaviors 130 (2022) 107310

Contents lists available at ScienceDirect

Addictive Behaviors

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

Psychometric evaluation of the NORC diagnostic screen for gambling problems (NODS) for the assessment of DSM-5 gambling disorder

Brad W. Brazeau *, David C. Hodgins

Department of Psychology, University of Calgary, Calgary, Canada

ARTICLE INFO

Keywords:
Gambling disorder
DSM-5
NODS
Validation
Psychometric
Assessment

ABSTRACT

The National Opinion Research Center (NORC) Diagnostic Screen for Gambling Problems (NODS) is one of the most used outcome measures in gambling intervention trials. However, a screen based on DSM-5 gambling disorder criteria has yet to be developed or validated since the DSM-5 release in 2013. This omission is possibly because the criteria for gambling disorder only underwent minor changes from DSM-IV to DSM-5: the diagnostic threshold was reduced from 5 to 4 criteria, and the illegal activity criterion was removed. Validation of a measure that captures these changes is still warranted. The current study examined the psychometric properties of an online self-report past-year adaptation of the NODS based on DSM-5 diagnostic criteria for gambling disorder (i.e., NODS-GD). A diverse sample of participants ($N = 959$) was crowdsourced via Amazon's TurkPrime. Internal consistency and one-week test–retest reliability were good. High correlations ($r = 0.74$–0.77) with other measures of gambling problem severity were observed in addition to moderate correlations ($r = 0.21$–0.36) with related but distinct constructs (e.g., gambling expenditures, time spent gambling, other addictive behaviors). All nine of the DSM-5 criteria loaded positively on one principal component, which accounted for 40% of the variance. Classification accuracy (i.e., sensitivity, specificity, predictive power) was generally very good with respect to the PGSI and ICD-10 diagnostic criteria. Future studies are encouraged to establish a gold standard self-report measure of gambling problems and develop agreed-upon recommendations for the use and interpretation of crowdsourced addiction data.

Humans have been gambling for millennia, from six-sided dice used five thousand years ago to contemporary online casinos. Nearly every culture has had a relationship with gambling, although societal acceptance has varied across time and context (Raylu & Oei, 2004). In the United States, gambling represents an occasional pastime for most of the adult population; however, past-year prevalence of gambling disorder (0.8%) remains high (Welte, Barnes, Tidwell, Hoffman, & Wieczorek, 2015). Given its simultaneous popularity and potential to cause significant problems (Hodgins, Stea, & Grant, 2011), it is critical that assessment instruments can accurately identify and classify varying levels of gambling engagement.

A screening instrument based on the original DSM-III criteria for pathological gambling became widely used in both clinical and community samples. More recently, the Problem Gambling Severity Index (PGSI; Ferris & Wynne, 2001) has gained prominence. The PGSI, which is not based on DSM diagnostic criteria, contains nine self-report items that measure gambling problem severity. Total scores are used to identify low-, moderate-, and high-risk gambling activity over the past year. The PGSI has demonstrated high correlations with other measures of gambling problem severity ($r = 0.83$), as well as good internal consistency ($\alpha = 0.86$), excellent specificity (1.0), and adequate sensitivity (0.83; Ferris & Wynne, 2001; Holtgraves, 2009).

The PGSI includes items that survey a range of problematic gambling characteristics, but does not offer comprehensive coverage of current DSM criteria nor yield a diagnosis. In contrast, the National Opinion Research Center (NORC) Diagnostic Screen for Gambling Problems (NODS; Gerstein et al., 1999) is a 17-item diagnostic interview that directly assesses gambling problems based on DSM-IV diagnostic criteria. The NODS correlates highly with other measures of gambling problem severity ($r = 0.86$) and moderately with log-transformed monthly gambling expenditures and number of days gambled ($r = 0.50$); it has also demonstrated fair internal consistency ($\alpha = 0.78$; Hodgins, 2004).

In 2013, the DSM-5 update included changes to the classification and diagnosis of gambling problems (American Psychiatric Association, 2013). Pathological gambling was reclassified as an addictive disorder

* Corresponding author at: Department of Psychology, University of Calgary, 2500 University Dr. NW, Calgary, AB T2N 1N4, Canada.
E-mail addresses: bradley.brazeau@ucalgary.ca (B.W. Brazeau), dhodgins@ucalgary.ca (D.C. Hodgins).

https://doi.org/10.1016/j.addbeh.2022.107310
Received 11 December 2021; Received in revised form 22 February 2022; Accepted 17 March 2022
Available online 18 March 2022
0306-4603/

given its significant overlap with substance use disorders in terms of comorbidity, symptom presentation, and biological characteristics (Petry, Blanco, Stinchfield, & Volberg, 2013). Consequently, it was renamed gambling disorder (GD) to reduce stigma and align with naming conventions of other addictive disorders. The illegal acts criterion was removed to reflect that it was often the last criterion endorsed by those diagnosed with DSM-IV pathological gambling; the diagnostic threshold was also reduced from 5 to 4 given the removal of a criterion. A DSM-5 diagnosis is typically derived by administering the NODS and excluding the illegal acts question; however, no measure has been developed or validated to reflect the 2013 DSM changes.

The purpose of the current study was to validate an online self-report past-year version of the NODS to assess DSM-5 GD via examination of psychometric properties (i.e., internal consistency, test–retest reliability, convergent and divergent validity, factor structure, item response patterns, sensitivity, specificity, and classification accuracy). Additionally, this version of the NODS was evaluated for how well it identifies ICD-10 pathological gambling.

1. Methods

1.1. Participants

Participants were recruited via Amazon's TurkPrime (Litman, Robinson, & Abberbock, 2017). TurkPrime is a virtual crowdsourcing platform that allows researchers to invite users to complete brief tasks (called Human Intelligence Tasks [HITs]) in exchange for financial compensation. To meet eligibility requirements, participants had to: a) be located in the United States; b) be 18 years of age or older; c) have gambled at least once in the past year; and d) demonstrate a HIT approval rate of 25% or greater (i.e., successful completion of at least 25% of attempted HITs).

Remuneration amounts on TurkPrime are typically based on the anticipated length of time it will take to complete a task. Best practices recommend a compensation rate of at least ten cents per minute (Chandler & Shapiro, 2016). For the current study, the anticipated completion time was ten minutes in total per worker; thus, participants were initially compensated a total of US$1.00. Mean survey completion times were greater than expected, which prompted the authors to increase worker remuneration to US$2.00. This study, including the methods, design, and modification, was approved by the University of Calgary Conjoint Faculties Research Ethics Board (REB20-1012).

2. Measures

The PGSI and NODS-GD were administered in self-report format using a past-year reporting window. A cut score of 8 or greater on the PGSI (i.e., problem gambling) was used to categorize participants; classifications were used as references for the measurement of NODS-GD classification accuracy. Note that there is no gold standard self-report measure of gambling problems; the PGSI was selected for the current study on the basis of its strong psychometric properties and frequent use in epidemiological research.

The Screener for Substance and Behavioral Addictions (SSBA) is a self-report measure that asks the same four questions as they apply to respondents' engagement with ten addictions over the past year, including gambling (Schluter, Hodgins, Wolfe, & Wild, 2018; Schluter, Kim, & Hodgins, 2018; Schluter, Hodgins, Thege, & Wild, 2020). Response options span from 0 (none of the time) to 4 (all of the time), and total scores range from 0 to 16 for each of the ten addiction scales. Each scale has demonstrated high internal consistency (α = 0.89–0.94) and moderate to high correlations with other measures of the same constructs (Schluter, Hodgins, Wolfe, & Wild, 2018; Schluter, Kim, & Hodgins, 2018).

The Composite International Diagnostic Interview gambling module (CIDI-GM) contains 17 yes/no questions that correspond to the four diagnostic criteria for pathological gambling in the International Statistical Classification of Diseases and Related Health Problems (ICD), tenth revision (World Health Organization, 2004). Participants that met all four criteria were classified as such as a reference point for measuring classification accuracy of the NODS for ICD-10 criteria. The CIDI-GM was administered in online self-report format along with the other measures included.

The Patient Health Questionnaire-9 (PHQ-9; Kroenke, Spitzer, & Williams, 2001) was included to assess discriminant validity. This self-report instrument contains nine questions that assess depressive symptoms over the past two weeks. Response options range from 0 (not at all) to 3 (nearly every day) and yield a total score between 0 and 27. The PHQ-9 has shown good internal consistency (α = 0.89; Kroenke et al., 2001). Participants who endorsed thoughts that they would be better off dead or hurting themselves were directed to an automated response at the end of the survey that encouraged participants to consult a resource (e.g., family physician) or contact a crisis helpline via phone numbers provided to them.

2.1. Procedure

Advertisements were displayed in TurkPrime to individuals that met eligibility criteria including IP addresses located in the United States. Eligible respondents were redirected to Qualtrics to complete the first part of the survey.

2.1.1. Part one: Screening

A virtual private network (VPN) block and reCAPTCHA system were implemented within part one to prevent the enrollment of ineligible and fake participants, respectively. Several demographic survey questions were then asked to crosscheck with TurkPrime filters and gather descriptive information from the sample. Participants were also asked to estimate the average number of hours they have gambled per month and per gambling session over the last three months, as well as the average net number of dollars they won or lost per month and per gambling session. These three-month retrospective self-report questions were adapted from the Gambling Participation Instrument (GPI; Williams, Volberg, Stevens, Williams, & Arthur, 2017).

Best practices recommend dual screening of participants recruited from platforms such as TurkPrime (Kim & Hodgins, 2017; Schluter, Hodgins, Wolfe, & Wild, 2018; Schluter, Kim, & Hodgins, 2018). To that end, two randomly selected PGSI questions were presented prior to the demographic questions, in addition to the full PGSI at the end of part one. Only individuals who met eligibility criteria and had matching PGSI responses were permitted to continue with the main survey (part two) immediately. Regardless, all participants who completed the first part were automatically compensated US$0.60.

2.1.2. Part two: Main survey

The main survey consisted of the NODS-GD, CIDI-GM, SSBA, and PHQ-9, in that order for all participants. Those who completed part two were manually compensated a bonus of US$0.60 and invited to complete the one-week follow-up.

2.1.3. Part three: One-week follow-up

The one-week follow-up comprised the readministered NODS-GD in addition to questions asking if participants were engaged and honest in their survey responses. Participants who completed part three were automatically credited with an additional US$0.80.

2.2. Data analysis

Data analysis covered three domains: reliability, validity, and classification accuracy. Test-retest reliability was assessed with the intraclass correlation coefficient (ICC) using a two-way mixed model to measure absolute agreement. Internal consistency was assessed with

Cronbach's alpha and McDonald's omega. Correlational analysis were used to evaluate convergent and divergent validity by calculating Pearson's correlation coefficients between the NODS-GD and measures of gambling problem severity, monthly gambling expenditures, hours per month spent gambling, depressive symptoms, and severity of concurrent addiction symptoms. The construct validity assessment entailed a principal components analysis using principal component extraction with promax rotation. Item response patterns on the PGSI and SSBA gambling scale were grouped by DSM-5 diagnostic severity (i.e., number of criteria met). Crosstabs were used to assess classification accuracy based on the PGSI and ICD-10 diagnostic criteria for pathological gambling. Sensitivity, specificity, positive predictive power, and negative predictive power were calculated from these crosstabs. Finally, percent agreement with the NODS-GD was calculated for the PGSI and CIDI-GM. All analyses were conducted with and without data from participants who did not endorse honest and attentive responses to survey questions; however, this did not significantly alter any results.

3. Results

All statistical analyses were conducted with R and RStudio (R Core Team, 2020). In total, 959 participants with a mean age of 39.4 (Mdn = 37, SD = 12.1) were recruited via TurkPrime from February to June 2021. Sample characteristics are provided in Tables 1 and 2. The one-week follow-up rate (50.8%) was much lower than would normally be expected from crowdsourcing platforms (Kim & Hodgins, 2017) despite attempts to increase retention (e.g., increased compensation, reduced HIT approval rate requirement) so follow-up invitations were discontinued for the final 315 participants.

3.1. Reliability

The mean NODS-GD score was 3.0 (SD = 2.5, Mdn = 2, mode = 1, range = 0–10). Of the 959 participants, 33% were low risk (score = 0–1), 33% were subthreshold (score = 2–3), 17% met criteria for mild GD (score = 4–5), 10% met criteria for moderate GD (score = 6–7), and 7% met criteria for severe GD (score = 8–9). Internal consistency of the NODS-GD (α = 0.88, ω = 0.89) was good, and one-week test–retest reliability (ICC = 0.84, p <.001) was also good (Cicchetti, 1994). Internal consistency of the PGSI in this sample was similar (α = 0.90, ω = 0.92). NODS-GD scores at follow-up (mean = 2.6, SD = 2.2, Mdn = 2) were slightly lower than baseline scores, $t(325)$ = 2.84, p =.005; this change is indicative of regression to the mean and is expected given that only a subsample was retested.

3.2. Validity

Table 3 displays the Pearson correlations between the NODS-GD, PGSI, and other measures. Convergent validity of the NODS-GD was strong, evidenced by high correlations with other direct measures of gambling severity (r = 0.74–0.77). As expected, moderate correlations were observed between the NODS-GD and measures of depression, other addictions, gambling expenditures, and hours spent gambling (r = 0.21–0.36), which are indicative of divergent validity. Item response patterns for direct measures are provided in Table 4 and broken down by NODS-GD classification.

The nine DSM-5 diagnostic criteria, based on the sixteen NODS-GD questions, were subjected to a principal components analysis to examine factor structure and establish construct validity. Results demonstrated a robust principal component that accounted for 40% of the variance and positively correlated with all nine criteria, which is indicative of a unitary construct. Two components had eigenvalues greater than one and together accounted for 51% of the variance. Promax rotated factor loadings for these two components are provided in Table 5. Five criteria loaded on Component A, which reflects cognitive-affective aspects of problematic gambling. In contrast, the four criteria that loaded on Component B reflect behavioral aspects.

Table 1
Sample demographics (N = 959).

Demographic variables	Descriptive statistics	
	n	%
Sex		
Male	503	52.6
Female	452	47.1
Non-binary	3	0.3
Education		
Some post-secondary or greater	751	78.3
Employment		
Employed full-time or part-time	703	73.3
Unemployed	76	7.9
Annual household income		
Under $30,000	172	17.9
$30,000 to $99,999	559	58.3
$100,000 and over	228	23.8
Relationship status		
Married or common-law	452	47.3
Single	383	40.0
Household members		
Live with spouse or common law partner	549	57.2
Live with child(ren) under 18	248	25.9
Live alone	196	20.4
Residential area		
Urban	621	64.8
Rural	323	33.7
Ethnicity		
White	757	78.9
Black	99	10.3
Asian American	72	7.5
Latin American	49	5.1
Native American	15	1.6
Religion		
Christianity	530	55.3
Atheism or agnosticism	352	36.7
Judaism	22	2.3
Buddhism	20	2.1
Islam	6	0.6

Note: Percentages do not always sum to 100 due to missing data and selective presentation of variables.

3.3. Classification accuracy

Classification accuracy of the NODS-GD based on DSM-5 and ICD-10 diagnostic criteria was evaluated with the PGSI and CIDI-GM categorizations, respectively (see Table 6). The NODS-GD classified more individuals as qualifying for GD (34%) compared to the PGSI (25%) and fewer compared to the CIDI-GM (40%). Percent agreement with the NODS-GD was 83% for the PGSI and 78% for the CIDI-GM. Sensitivity, specificity, and positive and negative predictive power of the NODS-GD are provided in Table 7.

4. Discussion

The results of this study support the use of a DSM-5 GD self-report adaptation of the NODS (i.e., NODS-GD) in a community sample of crowdsourced individuals who gamble; however, the PGSI may be

Table 2
Primary measures and descriptive statistics (N = 959).

Primary measures	Descriptive statistics		
	M	*Mdn*	*SD*
PGSI	6.03	4	6.53
NODS-GD	3.00	2	2.46
Hours spent gambling			
Per session	2.18	1	2.22
Per month	18.60	6	25.60
Dollars spent gambling			
Per session	104.21	25	269.11
Per month	560.28	100	1623.34
SSBA			
Gambling	3.14	2	3.77
Alcohol	2.60	1	3.58
Tobacco	2.70	0	4.65
Cannabis	1.62	0	3.40
Cocaine	0.47	0	1.94
Sex	1.82	0	3.26
Video gaming	2.69	1	3.59
Overeating	3.57	2	3.95
Overworking	2.42	0	3.63
Shopping	2.74	2	3.31
PHQ-9	6.49	5	5.77

Note: NODS-GD: National Opinion Research Center Diagnostic Screen for Gambling Problems, DSM-5 Gambling Disorder version; PGSI: Problem Gambling Severity Index; PHQ: Patient Health Questionnaire; SSBA: Screen for Substance and Behavioral Addictions.

Table 3
Pearson Correlations between Primary Measures and the NODS-GD and PGSI (N = 959).

Primary measures	NODS-GD	PGSI
PGSI	0.74	–
Hours spent gambling		
Per session	0.30	0.30
Per month	0.24	0.31
Dollars spent gambling		
Per session	0.23	0.24
Per month	0.21	0.29
SSBA		
Gambling	0.72	0.76
Shopping	0.36	0.35
Cocaine	0.30	0.40
Sex	0.30	0.35
Video gaming	0.29	0.25
Tobacco	0.28	0.33
Overworking	0.28	0.27
Overeating	0.27	0.25
Cannabis	0.26	0.27
Alcohol	0.24	0.30
PHQ-9	0.32	0.33

Note: All $p < .001$.
NODS-GD: National Opinion Research Center Diagnostic Screen for Gambling Problems, DSM-5 Gambling Disorder version; PGSI: Problem Gambling Severity Index; PHQ: Patient Health Questionnaire; SSBA: Screen for Substance and Behavioral Addictions.

preferrable if researchers are not interested in using diagnostic classifications. Internal consistency and one-week test–retest reliability of the NODS-GD were both good. High correlations were observed between the NODS-GD and other direct measures of gambling problem severity (i.e., PGSI, SSBA gambling scale); moderate correlations were observed between the NODS-GD and other validation measures (i.e., gambling expenditures, time spent gambling, depressive symptoms, SSBA non-gambling scales). Correlations with primary measures were generally the same or stronger for the PGSI compared to the NODS-GD. While one might predict that gambling hours and expenditures would be more highly correlated with the NODS-GD, moderate correlations are to be expected given the different timeframes (past year versus past three months) and distinct constructs (gambling behaviors versus diagnostic problem severity).

All nine of the DSM-5 criteria loaded positively on the principal component, which accounted for 40% of the variance; this suggests that the NODS-GD measures a single construct. Two components with eigenvalues greater than one roughly represented cognitive and behavioral symptoms of GD, which is consistent with a cognitive-behavioral theoretical formulation of gambling problems. Classification accuracy of the NODS-GD was generally very good with reference to the PGSI and CIDI-GM. The NODS-GD classified more individuals as GD compared to the PGSI and fewer compared to the CIDI-GM. Importantly, the lack of a gold standard self-report measure of gambling problem severity obfuscates possible conclusions.

The findings of this paper should be interpreted in the context of limitations inherent to crowdsourced data. Participants in the current study were self-selected, and those who demonstrated eligibility were financially compensated. Thus, there may have been an incentive for prospective participants to misrepresent themselves in order to obtain compensation. TurkPrime samples also generally report higher rates of gambling problems compared to the general population (Schluter, Hodgins, Wolfe, & Wild, 2018; Schluter, Kim, & Hodgins, 2018), which may in part be indicative of malingering (Pickering & Blaszczynski, 2021). However, some researchers argue that the anonymous nature of TurkPrime may actually reduce bias associated with self-report of stigmatizing psychopathology (Russell, Browne, Hing, Rockloff, & Newall, 2021). Our HIT approval rate cutoff for participants was also quite low (25%) compared to the recommended cutoff of 95% (Kim & Hodgins, 2017; Schluter, Hodgins, Wolfe, & Wild, 2018; Schluter, Kim, & Hodgins, 2018); low approval rates imply that TurkPrime users have provided unusable data to a large proportion of prior HITs. Unfortunately, users with extremely high approval rates are much less common; the use of a low approval rate was necessary to ensure timely recruitment of participants.

We employed several strategies and best practices to bolster data quality. Remuneration amounts adhered to ethical standards (Chandler & Shapiro, 2016), which aim to balance adequate compensation with minimal incentive for participants to misrepresent themselves. Filters in TurkPrime and Qualtrics were used to prevent users from completing the study more than once; in one study, this strategy alone reduced fraudulent responses by 80% (Chandler & Paolacci, 2017). IP addresses were screened by Qualtrics to ensure participants were located in the US, and those concealing their location with a VPN were also excluded. A reCAPTCHA screener blocked internet bots from submitting any responses. Participants with non-matching responses to two randomly selected PGSI questions were also excluded; this dual-screening strategy served as both an attention check and a method to limit dishonest responses. Taken together, these strategies represent several recommendations for enhancing data quality proposed by both critics and proponents of crowdsourced convenience samples (Chandler & Shapiro, 2016; Kim & Hodgins, 2017; Pickering & Blaszczynski, 2021). It does appear that negative consequences have been mitigated; for example, filtering out participants who admitted after survey completion to responding dishonestly or inattentively did not significantly alter any findings. The data were certainly adequate given the purpose of the current study, although TurkPrime may be less suitable for other research objectives (e.g., estimating population prevalence). Regardless, future research would benefit from implementing more of these strategies (e.g., multiple attention checks, pre-registration) when crowdsourcing or using alternative platforms explicitly designed for research (e.g., Prolific).

Table 4
Mean Scores on Measures of Gambling Problems by NODS-GD Classification (N = 959).

Measures	Classification				
	Low Risk (0–1)	Subthreshold (2–3)	Mild GD (4–5)	Moderate GD (6–7)	Severe GD (8–9)
PGSI	2.08	4.13	7.51	14.51	18.00
SSBA: Gambling	0.92	2.01	4.31	7.37	10.04

Note: GD: gambling disorder; NODS-GD: National Opinion Research Center Diagnostic Screen for Gambling Problems, DSM-5 Gambling Disorder version; PGSI: Problem Gambling Severity Index; PHQ: SSBA: Screen for Substance and Behavioral Addictions.

Table 5
Promax Rotated Factor Loadings for the NODS-GD (N = 959).

DSM-5 criterion	Component	
	A	B
Loss of control	0.38	
Dishonesty	0.35	
Preoccupation	0.33	
Tolerance	0.33	
Escape from distress	0.27	
Financial bailout		0.57
Life problems		0.52
Failed quit attempts		0.42
Chasing losses		0.31

Note: NODS-GD: National Opinion Research Center Diagnostic Screen for Gambling Problems, DSM-5 Gambling Disorder version; PGSI: Problem Gambling Severity Index.

Table 6
Classification Accuracy of the NODS-GD (N = 959).

NODS-GD classification	PGSI classification		Total	CIDI classification		Total
	No	Yes		No	Yes	
No (does not meet criteria)	595 (62%)	38 (4%)	633 (66%)	503 (52%)	131 (14%)	634 (66%)
Yes (does meet criteria)	121 (13%)	205 (21%)	326 (34%)	79 (8%)	246 (26%)	325 (34%)
Total	716 (75%)	243 (25%)		582 (60%)	377 (40%)	

Note: CIDI: Composite International Diagnostic Interview; GD: gambling disorder; NODS-GD: National Opinion Research Center Diagnostic Screen for Gambling Problems, DSM-5 Gambling Disorder version.

Table 7
Sensitivity, Specificity, and Predictive Power of the NODS-GD (N = 959).

NODS-GD measurement	Comparison	
	PGSI	CIDI-GM
Sensitivity	84%	65%
Specificity	83%	86%
Positive predictive power	63%	76%
Negative predictive power	94%	79%

Note: CIDI-GM: Composite International Diagnostic Interview, gambling module; NODS-GD: National Opinion Research Center Diagnostic Screen for Gambling Problems, DSM-5 Gambling Disorder version. PGSI: Problem Gambling Severity Index.

5. Conclusion

The NODS-GD represents a reliable and valid measure of gambling problem severity, based on DSM-5 diagnostic criteria, when administered in online self-report format to a diverse crowdsourced sample of community gamblers. Future validation studies are encouraged to establish a gold standard self-report measure of gambling problem severity for the psychometric evaluation of classification accuracy. Given the several benefits (e.g., cost-effective recruitment of diverse samples) and drawbacks (e.g., data quality concerns) of crowdsourced data, the field would also benefit from agreed-upon recommendations for interpretation of such data in addiction populations.

Role of Funding Sources

This work was supported by a student research allowance from the Alberta Gambling Research Institute (AGRI). AGRI had no influence over study design, data collection, data analysis, interpretation, manuscript writing, or manuscript submission.

CRediT authorship contribution statement

Brad W. Brazeau: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Funding acquisition. **David C. Hodgins:** Conceptualization, Methodology, Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th edn.). Washington, DC: American Psychiatric Publishing Inc.
Chandler, J., & Paolacci, G. (2017). Lie for a dime: When most prescreening responses are honest but most study participants are imposters. *Social Psychological and Personality Science, 8*(5), 500–508.
Chandler, J., & Shapiro, D. (2016). Conducting clinical research using crowdsourced convenience samples. *Annual Review of Clinical Psychology, 12*(1), 53–81.
Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6*(4), 284–290.
Ferris, J., & Wynne, H. (2001). *The Canadian Problem Gambling Index: Final report.* Ottawa, ON: Canadian Centre on Substance Abuse.
Gerstein, D., Hoffmann, J., Larison, C., Engelman, L., Murphy, S., & Palmer, A. (1999). Gambling impact and behavior study. *National Gambling Impact Study Commission Report*, Exec-2.
Hodgins, D. C. (2004). Using the NORC DSM Screen for Gambling Problems as an outcome measure for pathological gambling: Psychometric evaluation. *Addictive Behaviors, 29*(8), 1685–1690.
Hodgins, D. C., Stea, J. N., & Grant, J. E. (2011). Gambling disorders. *The Lancet, 378*(9806), 1874–1884.
Holtgraves, T. (2009). Evaluating the problem gambling severity index. *Journal of Gambling Studies, 25*(1), 105–120.
Kim, H. S., & Hodgins, D. C. (2017). Reliability and validity of data obtained from alcohol, cannabis, and gambling populations on Amazon's Mechanical Turk. *Psychology of Addictive Behaviors, 31*(1), 85–94.
Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General International Medicine, 16*(02), 606–613.
Litman, L., Robinson, J., & Abberbock, T. (2017). TurkPrime.com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior Research Methods, 49*(2), 433–442.
Petry, N. M., Blanco, C., Stinchfield, R., & Volberg, R. (2013). An empirical evaluation of proposed changes for gambling diagnosis in the DSM-5. *Addiction, 108*(3), 575–581.
Pickering, D., & Blaszczynski, A. (2021). Paid online convenience samples in gambling studies: Questionable data quality. *International Gambling Studies, 21*(3), 516–536.

R Core Team. (2020). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. http://www.R-project.org.
Raylu, N., & Oei, T. P. (2004). Role of culture in gambling and problem gambling. *Clinical Psychology Review, 23*(8), 1087–1114.
Russell, A. M. T., Browne, M., Hing, N., Rockloff, M., & Newall, P. (2021). Are any samples representative or unbiased? Reply to Pickering and Blasczcynski. *International Gambling Studies*, 1–12.
Schluter, M. G., Hodgins, D. C., Thege, B. K., & Wild, T. C. (2020). Predictive utility of the brief Screener for Substance and Behavioral Addictions for identifying self-attributed problems. *Journal of Behavioral Addictions, 9*(3), 709–722.
Schluter, M. G., Hodgins, D. C., Wolfe, J., & Wild, T. C. (2018). Can one simple questionnaire assess substance-related and behavioral addiction problems? Results of a proposed new screener for community epidemiology. *Addiction, 113*(8), 1528–1537.
Schluter, M. G., Kim, H. S., & Hodgins, D. C. (2018). Obtaining quality data using behavioral measures of impulsivity in gambling research with Amazon's Mechanical Turk. *Journal of Behavioral Addictions, 7*(4), 1122–1131.
Welte, J. W., Barnes, G. M., Tidwell, M.-C.-O., Hoffman, J. H., & Wieczorek, W. F. (2015). Gambling and problem gambling in the United States: Changes between 1999 and 2013. *Journal of Gambling Studies, 31*(3), 695–715.
Williams, R. J., Volberg, R. A., Stevens, R. M. G., Williams, L. A., & Arthur, J. N. (2017). The definition, dimensionalization, and assessment of gambling participation. *Report Prepared for the Canadian Consortium for Gambling Research*, 1–157.
World Health Organization. (2004). *ICD-10: International statistical classification of diseases and related health problems* ((10th ed.).). World Health Organization.