International Gambling Studies

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/rigs20

Assessing severity of problem gambling – confirmatory factor and Rasch analysis of three gambling measures

Olof Molander & Peter Wennberg

INTERNATIONAL GAMBLING STUDIES
https://doi.org/10.1080/14459795.2022.2149834

Routledge
Taylor & Francis Group

RESEARCH ARTICLE

OPEN ACCESS Check for updates

Assessing severity of problem gambling – confirmatory factor and Rasch analysis of three gambling measures

Olof Molander [a] and Peter Wennberg [b,c,d]

[a]Department of Clinical Neuroscience, Center for Psychiatry Research, Karolinska Institutet, Solna, Sweden; [b]Department of Public Health Sciences, Stockholm University, Stockholm, Sweden; [c]Karolinska Institutet, Department of Global Public Health, Solna, Sweden; [d]Inland Norway University of Applied Sciences, Department of Psychology

ABSTRACT

The comparative psychometric properties of self-report measures for gambling are insufficiently evaluated, in particular regarding factor structure and item response properties. Confirmatory factor and Rasch analyses were tested for three widely used gambling measures assessing problem gambling and related constructs, that is, the Problem Gambling Severity Index (PGSI), the Problem and Pathological Gambling Measure (PPGM), and the NORC Diagnostic Screen for Gambling Problems (NODS). Psychometric data was analyzed, including help-seeking and recreational gambling samples (N = 598). Compared to the PPGM and the NODS, the PGSI performed worse in the confirmatory factor analysis, and showed poor fit for the theoretically assumed unidimensional model. The Rasch analysis indicated that the PPGM had an adequate difficulty range (i.e. lowest to highest item difficulty) to detect gambling problems across a severity continuum. Compared to the PPGM, the PGSI and NODS had smaller item difficulty ranges, indicating detection of higher gambling severity problems. We conclude that using the PGSI for detection of low severity problems, such as at-risk gambling, might be problematic. The PPGM can be used in general populations and clinical contexts to detect problem gambling and pathological gambling. The NODS is suitable for use in clinical samples for identification of pathological gambling.

ARTICLE HISTORY
Received 9 March 2022
Accepted 13 November 2022

KEYWORDS
The problem gambling severity index; the problem and pathological gambling measure; the NORC diagnostic screen for gambling problems; confirmatory factor analysis; Rasch analysis

Introduction

A large proportion of gambling research has focused on the public health concept of problem gambling, defined, for example as ‘excessive gambling behavior that creates negative consequences for the gambler, others in his/her social network, and for the community’ (Blaszczynski & Nower, 2002; Molander et al., 2021). In the 5th edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5; American Psychiatric Association, 2013), the previous clinical criteria pathological gambling (American Psychiatric Association, 1994) was replaced with Gambling Disorder. The current criteria includes three diagnostic symptom severity levels, i.e., mild, moderate or

CONTACT Olof Molander olof.molander@ki.se Department of Clinical Neuroscience, Center for Psychiatry Research, Karolinska Institutet, Solna, Sweden

severe Gambling Disorder. Another change in the DSM-5 diagnostic criteria was that the criterion illegal acts to finance gambling was removed. Compared to pathological gambling, relatively less is known of Gambling Disorder. The shift to a different diagnostic category and changes to the criteria and specifications has complicated accurate/reliable measurement of this clinical construct using existing gambling instruments (Molander et al., 2021; Otto et al., 2020).

Estimates of past year prevalence of problem gambling across countries have ranged between 0.12% and 5.8% (see Calado & Griffiths, 2016 for a review). Several measurement issues have been noted which have affected problem gambling prevalence estimates and rendered comparisons between studies difficult, for example, variability in measures used to assess problem gambling, differences in problem gambling scoring thresholds used for the same gambling measure, various time frames used to assess problem gambling, and variations in item content (Molander et al., 2019; Williams et al., 2012).

Three widely used gambling measures to assess problem gambling are the Problem Gambling Severity Index (PGSI; Ferris & Wynne, 2001), the Problem and Pathological Gambling Measure (PPGM; Williams & Volberg, 2013) and the NORC Diagnostic Screen for Gambling Problems (NODS; Gerstein et al., 1999). Briefly, the PGSI was developed from the Canadian Problem Gambling Index (CPGI) as a measure to assess at-risk and problem gambling in population surveys (Ferris & Wynne, 2001). The PPGM was developed as a comprehensive measure to assess all gambling-related harms. The NODS was developed as a DSM-IV-based measure to assess higher clinical severity. Compared to the PGSI, the PPGM and the NODS also includes scoring for pathological gambling according to DSM-IV to facilitate use in clinical samples as well as the general population (PGSI cutoffs have also been established in clinical samples; see for example Merkouris et al., 2020). Table 1 shows an overview of psychometric evaluations of the PGSI, the PPGM and the NODS in their respective original studies.

Although the PGSI, the PPGM and the NODS have shown satisfactory psychometric properties such as internal consistency, convergent and divergent validity (see, for example, Hodgins, 2004; Holtgraves, 2009; Wickwire et al., 2008), other psychometric estimates have resulted in mixed results or remain to be investigated. First, the PGSI, the PPGM and the NODS have mainly been evaluated in separate samples which complicates relevant comparisons. Second, most studies have used explorative factor analysis, and not confirmatory factor analysis (CFA), which might indicate a widespread uncertainty regarding the theoretical factor structures of the instruments. In terms of factor structure, both the PGSI and the NODS have been assumed to have a unifactorial structure which includes items within a single theoretical domain. Overall, previous psychometric studies have supported this but possible multi-dimensionality or presence of subfactors has also been suggested for both PGSI and the NODS (Christensen et al., 2019; Ferris & Wynne, 2001; Hodgins, 2004; Holtgraves, 2009; Toce-Gerstein et al., 2003). The PPGM is assumed to have a three-factor structure encompassing items within the theoretical domains Problems, Impaired control, and Other Issues. The factor structure of the PPGM has not yet been statistically evaluated according to our knowledge (personal communication Rachel Volberg, 1 July 2021). Third, most psychometric evaluations have been based on classical test theory. In contrast to classical test theory which evaluate reliability and validity of measures based on their items, item response theory approaches estimate items and persons on a continuum and defines the relative positions of these on

Table 1. Overview of psychometric original studies.

					Original study		
Measure	Development aims and intended use	*N* items	Sub-scales (factors)	Scoring classification	Psychometric evaluation[a]	Sample	Reference
PGSI	'[A measure for] use in general population surveys, one that reflected a more holistic view of gambling, and included more indicators of social context.' Ferris and Wynne (2001)	9	No, only total score.	No problems, low at-risk, moderate at-risk, problem gambling	Cronbachs α = 0.84. Test-retest reliability = 0.78. Factor analysis supported unidimensionality. Sensitivity 78% and 83%, specificity 100% and 100%, compared to clinical assessment interviews and DSM-IV, respectively.	General population	Ferris and Wynne (2001)
PPGM	'Utility in both population surveys and clinical settings' "[Assessment of] all potential harms deriving from gambling" Williams and Volberg (2013)	14	Yes: Problems, Loss of control, and Other + Total score		Cronbachs α = 0.81. Test-retest reliability = 0.78. Concurrent validity Kendall's Tau = 0.69–0.78 to other gambling measures. Diagnostic efficiency 99%.	General population	Williams and Volberg (2013)
NODS	'[A measure] designed to be more demanding and restrictive in assessing problematic behaviors than (. . .) other screens based on the DSM – IV criteria' Gerstein et al. (1999)	17	No, only total score.	No problems, mild/subclinical risk, moderate/subclinical risk, pathological gambling[b]	Test-retest reliability, *r* = 0.99 and 0.98, for lifetime and past-year, respectively.	Outpatient treatment sample. General population.	Gerstein et al. (1999)

[a] = Key results summarized, [b] = The NODS scoring classification was later revised to No problems, At-risk gambling, Problem gambling, and Pathological gambling (Wickwire et al., 2008).
NODS = The NORC Diagnostic Screen for Gambling Problems (Gerstein et al., 1999), 30 days version.
PGSI = The Problem Gambling Severity Index (Ferris & Wynne, 2001).
PPGM = The Problem and Pathological Gambling Measure (PPGM; Williams & Volberg, 2013).

a severity scale (Cappelleri et al., 2014; Cowlishaw et al., 2019; Wilson, 2004). Advantages of item response theory models, such as Rasch analysis, includes item-specific evaluations of fit and difficulty (severity), as well as establishment of measures' item difficulty range across a problem gambling severity continuum. Furthermore, if a measure shows good fit in a Rasch model, parametric statistical methods can be used on an interval level. A few studies have investigated the PGSI, the PPGM and the NODS using Rasch analysis (Cowlishaw et al., 2019; Miller et al., 2013; Molde et al., 2010). These studies have indicated that most individual items show good fit with varying difficulty, but also that the PGSI, the PPGM and the NODS mainly are suited for discriminating across more severe levels of problem gambling which could be problematic especially for the PGSI.

The PGSI, the PPGM, and the NODS are all widely used to assess problem gambling and related constructs. However, from a clinical perspective it is unclear which measure is optimal for assessment of gambling severity in a certain population. From a research perspective, it is important to establish the measures' metric properties to enable use in advanced statistical analyses. Ideally, a Rasch analysis could establish that an instrument lies on a data level that allows parametric analyses. Consequently, the aim of the current study was to test the factor structure and Rasch estimates of the PGSI, the PPGM, and the NODS among gamblers from recreational and help-seeking samples.

Material and methods

Sample and procedure

Existing data from a previous psychometric evaluation of a novel gambling measure was analyzed (Molander et al., 2021). Briefly, the study by Molander et al. (2021) recruited Swedish gamblers from four different samples, that is, recreational, support and treatment seeking gamblers, and gamblers from self-help groups, through advertisements and clinicians within the healthcare. The participants completed self-report measures in an online survey which also included the PGSI, the PPGM and the NODS ($n = 598$ were assessed with these measures). The study was approved by the Regional Ethics Board of Stockholm, Sweden (ref. no. 2017/1479–31), and all participants provided informed consent for study participation and publication of results. In the current study, the support and treatment-seeking gamblers, as well as the gamblers from self-help groups, were collapsed into a help-seeking cohort ($n = 306$). See Table 2 for demographic characteristics.

Measures

The PGSI

The PGSI (Ferris & Wynne, 2001) is a 9-item self-report measure assessing at-risk and problem gambling during the past year. The PGSI uses multiple response alternatives, that is, 'Never', 'Sometimes', 'Most of the time', and 'Always'. The PGSI scoring index (Total score max 27) is classified into four levels of gambling: No gambling problems (Total score 0), low risk level of problems with few or no identified negative consequences (Total score 1–2), moderate risk level of problems leading to some negative consequences (Total score 3–7), and problem gambling with negative consequences and a possible loss of control (Total score ≥8).

The PPGM

The PPGM (Williams & Volberg, 2013) is a self-report measure assessing gambling-related negative consequences during the past year. PPGM consist of 14 items, including four sub-items. The PPGM scoring index collapses items 1a and 1b, items 3a and 3b, and items 5a and 5b, into three items, respectively, which are scored 0 or 1. Item 10a is not scored; item 10b is scored as item 10. PPGM uses with dichotomous (yes or no) response alternatives with a Total score (max 14) and three sub-scales: Problems (items 1–7), Impaired control (items 8–11) and Other Issues (items 12–14). The PPGM scoring index uses a diagnostic algorithm resulting in four levels of gambling: recreational, at-risk, problem gambler and pathological gambling.

The NODS

The NODS (Gerstein et al., 1999; Wickwire et al., 2008) is a 17-item self-report measure assessing no problems, at-risk, problem gambling and pathological gambling. The NODS uses dichotomous (yes or no) response alternatives. Three different time-based NODS versions are available (i.e. lifetime, past year, and past 30 days) and in the current study the NODS 30 days version was used. In the NODS scoring index items 1 and 2, items 8 and 9, and items 14, 15 and 16, are collapsed into three items, respectively, which are scored 0 or 1. Items 4, 6, and 11 are not scored. The NODS scoring (Total score max 10) is classified into four levels of gambling: no problems (Total score 0), at-risk (Total score 1–2), problem gambling (Total score 3–4), and pathological gambling (Total score ≥5) (Wickwire et al., 2008).

Demographic characteristics

The online survey included a set of items assessing demographic and gambling characteristics, for instance age, sex, education, civil status, and gambling types (see Table 2).

Statistical analyses

Due to an administration error in the data collection, 167 participants had missing data for $\text{NODS}_{\text{item10}}$ and one participant for $\text{NODS}_{\text{item6}}$. These missing data were replaced using multiple imputations by chained equations, a technique with valid statistical inference which restores the natural variability of the missing values and addresses the uncertainty due to the missing data (Enders et al., 2016; Kang, 2013). Analyses were performed on scoring-relevant items, of which some were collapsed according to the scoring procedure of the PPGM (max item score = 1, see the PPGM above). CFA models were used to test proposed theoretical domains for the PGSI (unifactorial structure), the PPGM (three-factor and unifactorial structure), and the NODS (unifactorial structure). The CFA used a maximum likelihood estimator for the PGSI, and diagonally weighted least squares for the PPGM and the NODS due to use of dichotomous yes/no response alternatives. The CFA analyses included model fit indices (see Bowen & Guo, 2011 for fit indices cutoffs), as well as standardized item factor loadings. In addition to the PGSI CFA, a follow-up exploratory factor analysis was performed (see Factor analysis below). Cronbachs alpha (α) estimates of measures' total scores were used to test internal

Table 2. Demographic characteristics across gambling samples.

	Gambling population		
	Recreational (n=292)	Help-seeking (n=306)	Total (N=598)
Demographic characteristics			
Age M (Sd)	36.4 (12.7)	29.5 (10.5)	33 (12.1)
Sex (%)			
Men	66	82	74
Women	33	17	25
Not stated	1	1	1
Source of income (%)			
Employed	68	60	64
Studies	12	30	21
Other[a]	20	10	15
Highest level of education (%)			
University	25	46	35
High school	60	45	52
Junior high school	13	8	10
Civil status (%)			
Cohabiting	51	54	52
Children	51	30	41
Gambling characteristics:			
Gambling debts (%)	63	8	36
Gambling types (%)[b]			
Casino online	67	33	51
Casino land-based	12	12	12
Sport games online	35	53	44
Sport games venue	15	13	14
Poker online	14	20	17
Poker club	5	6	6
EGM	10	4	7
Number games	9	9	9
Lotteries	18	35	27
Horse betting	18	16	17
Bingo	11	9	10
Other	8	14	11

[a]This category included unemployment insurance, income support, sickness compensation, sickness benefit, pension, and other sources of income.
[b]Participants were able to report several gambling types.

consistency in the total sample and among sub-groups (see Table 5). Rasch analyses were used to test item difficulty for specific items across a severity continuum, as well as measures' item difficulty ranges. Infit and outfit are item specific Rasch fit indices which indicates how accurately or predictably the data fits the model. Infit and outfit estimates >1.50 indicates item underfit and might be unproductive for measurement. Infit and outfit estimates <0.50 indicates item overfit and can be less productive for measurement. Estimates within the range 0.5 to 1.5 are acceptable and productive for measurement (Miller et al., 2013; Wright & Linacre, 1994; Linacre, 2002). Person reliability tests whether a measure can discriminate the sample into enough levels (or strata) given the purpose of the measure (cutoffs 0.5 = 1 or 2 levels, 0.8 = 2 or 3 levels, 0.9 = 3 or 4 levels; Bond & Fox, 2007). Dichotomous Rasch models were used for all measures. For the PGSI, an additional Rasch rating-scale model was used as the PGSI uses multiple response alternatives (Wind & Hua, 2021). Analyses were performed using R Studio and Jamovi (R Core Team, 2018) with the following key packages: mice, foreign, lavaan, psych and TAM. See Table 3 for measure scores.

Table 3. Total scores of PGSI, PPGM and NODS.

Total score, frequency (%)	Recreational (n=292)			Help-seeking (n=306)			Total (N=598)		
	PGSI	PPGM	NODS	PGSI	PPGM	NODS	PGSI	PPGM	NODS
0	136 (46.6%)	171 (58.6%)	206 (70.5%)	54 (17.6%)	64 (20.9%)	110 (35.9%)	190 (31.8%)	235 (39.3%)	316 (52.8%)
1	27 (9.2%)	36 (12.3%)	30 (10.3%)	11 (3.6%)	13 (4.2%)	38 (12.4%)	38 (6.4%)	49 (8.2%)	68 (11.4%)
2	32 (11%)	17 (5.8%)	15 (5.1%)	13 (4.2%)	9 (2.9%)	18 (5.9%)	45 (7.5%)	26 (4.3%)	33 (5.5%)
3	18 (6.2%)	10 (3.4%)	10 (3.4%)	3 (1%)	9 (2.9%)	18 (5.9%)	21 (3.5%)	19 (3.2%)	28 (4.7%)
4	11 (3.8%)	9 (3.1%)	7 (2.4%)	3 (1%)	6 (2%)	23 (7.5%)	14 (2.3%)	15 (2.5%)	30 (5%)
5	10 (3.4%)	11 (3.8%)	8 (2.7%)	4 (1.3%)	11 (3.6%)	17 (5.6%)	14 (2.3%)	22 (3.7%)	25 (4.2%)
6	8 (2.7%)	9 (3.1%)	8 (2.7%)	6 (2%)	9 (2.9%)	23 (7.5%)	14 (2.3%)	18 (3%)	31 (5.2%)
7	4 (1.4%)	5 (1.7%)	5 (1.7%)	8 (2.6%)	23 (7.5%)	20 (6.5%)	12 (2%)	28 (4.7%)	25 (4.2%)
8	5 (1.7%)	4 (1.4%)	1 (0.3%)	4 (1.3%)	19 (6.2%)	17 (5.6%)	9 (1.5%)	23 (3.8%)	18 (3%)
9	7 (2.4%)	6 (2.1%)	2 (0.7%)	9 (2.9%)	22 (7.2%)	11 (3.6%)	16 (2.7%)	28 (4.7%)	13 (2.2%)
10	3 (1%)	6 (2.1%)	-	7 (2.3%)	31 (10.1%)	11 (3.6%)	10 (1.7%)	37 (6.2%)	11 (1.8%)
11	1 (0.3%)	5 (1.7%)		8 (2.6%)	35 (11.4%)		9 (1.5%)	40 (6.7%)	
12	3 (1%)	2 (0.7%)		9 (2.9%)	24 (7.8%)		12 (2%)	26 (4.3%)	
13	4 (1.4%)	1 (0.3%)		12 (3.9%)	23 (7.5%)		16 (2.7%)	24 (4%)	
14	2 (0.7%)	-		8 (2.6%)	8 (2.6%)		10 (1.7%)	8 (1.3%)	
15	3 (1%)			8 (2.6%)			11 (1.8%)		
16	2 (0.7%)			18 (5.9%)			20 (3.3%)		
17	4 (1.4%)			14 (4.6%)			18 (3%)		
18	3 (1%)			8 (2.6%)			11 (1.8%)		
19	5 (1.7%)			13 (4.2%)			18 (3%)		
21	1 (0.3%)			12 (3.9%)			13 (2.2%)		
22	3 (1%)			11 (3.6%)			14 (2.3%)		
20	-			10 (3.3%)			10 (1.7%)		
23	-			9 (2.9%)			9 (1.5%)		
24	-			15 (4.9%)			15 (2.5%)		
25	-			13 (4.2%)			13 (2.2%)		
26	-			8 (2.6%)			8 (1.3%)		
27	-			8 (2.6%)			8 (1.3%)		

NODS = The NORC Diagnostic Screen for Gambling Problems (Gerstein et al., 1999), 30 days version. PGSI = The Problem Gambling Severity Index (Ferris & Wynne, 2001). PPGM = The Problem and Pathological Gambling Measure (PPGM; Williams & Volberg, 2013).

Results

Factor analyses

CFA were estimated for the PGSI and the NODS, using unidimensional factor models. All standardized factor loadings were acceptable but fit indices were poor for the PGSI (see Table 4). To illustrate this further, a follow-up exploratory factor analysis (principal axis factoring extraction method with varimax rotation) was conducted for the PGSI. This indicated that a three-factor solution showed best fit for the data (RMSEA = 0.036, TLI = 0.995, DF = 12, $\chi2$ = 21.5) where factor 1 (PGSI items 1, 3, 5–9, had factors loadings >0.40), factor 2 (PGSI items 1–5, 7) and factor 3 (PGSI items 4, 8) accounted for 36%, 29% and 13% of the variance, respectively. Regarding the PPGM, we estimated two CFA models; one three-factor model consisting of the proposed PPGM sub-scales (Problem, Impaired control, and Other Issues), and one single-factor model. The results (see Table 4) indicated that both models showed satisfactory fit and that all standardized loadings were acceptable except for $PPGM_{item6}$ (gambling-related illegal acts).

Table 4. Confirmatory factor analyses for PGSI, PPGM and NODS (N = 598).

			Fit indices					Standardized item loadings on CFA	
Measure	Factor structure	Estimator	RMSEA (90% CI)	CFI	TLI	DF	χ2	Range	Items <0.40
PGSI	One factor model	ML	0.135 (0.122–0.148)	0.948	0.931	27	320.1	0.68–0.92	-
PPGM	Three factor model	DWLS	0.030 (0.019–0.041)	1.00	1.00	74	114.7	0.36–0.91	$PPGM_{item6}$
	One factor model	DWLS	0.042 (0.033–0.051)	1.00	0.96	77	158.8	0.34–0.88	$PPGM_{item6}$
NODS	One factor model	DWLS	0.000 (0.000–0.021)	1.00	1.00	35	28.5	0.53–0.81	-

Diagonally weighted least squares were used for PPGM and NODS, as these measures uses dichotomous response alternatives.
CFA = Confirmatory factor analysis.
CFI = Comparative Fit Index.
DWLS = Diagonally weighted least squares.
ML = Maximum likelihood.
NODS = The NORC Diagnostic Screen for Gambling Problems (Gerstein et al., 1999), 30 days version.
PGSI = The Problem Gambling Severity Index (Ferris & Wynne, 2001).
PPGM = The Problem and Pathological Gambling Measure (PPGM; Williams & Volberg, 2013).
RMSEA = Root Mean Square Error of Approximation.
TLI = Tucker-Lewis Index.

Internal consistency

Cronbachs α was estimated for the total sample, as well as per sex, age and gambling groups. Initially, internal consistency was estimated across five age groups (18–30, 31–40, 41–50, 51–60 and >60 years). However, as the age group >60 years was small (n = 18) and showed lack of variance for several items. Therefore, the initial age groups 51–60 and >60 years were collapsed into one group, >51 years. The results indicated good to excellent internal consistency for all measures and showed small differences across sub-groups (see Table 5).

Table 5. Internal consistency for PGSI, PPGM and NODS, across sex, age and gambling groups.

		Sex		Age (years)				Gambling population	
Cronbachs α[a] (95% CI)	Total N = 598	Men n = 441	Women n = 150	18-30 n = 304	31-40 n = 146	41-50 (n = 86)	>51 (n = 60)	Recreational (n = 292)	Help-seeking (n = 306)
PGSI	0.96 (0.95–0.96)	0.96 (0.95–0.96)	0.96 (0.96–0.97)	0.95 (0.95–0.96)	0.97 (0.96–0.98)	0.96 (0.94–0.97)	0.95 (0.94–0.97)	0.93 (0.92–0.94)	0.95 (0.94–0.96)
PPGM[c]	0.94 (0.93–0.95)	0.94 (0.93–0.94)	0.95 (0.94–0.96)	0.93 (0.92–0.94)	0.95 (0.94–0.96)	0.94 (0.92–0.96)	0.93 (0.91–0.96)	0.90 (0.89–0.92)	0.92 (0.91–0.94)
NODS	0.90 (0.88–0.91)	0.88 (0.87–0.90)	0.91 (0.89–0.93)	0.87 (0.85–0.89)	0.94 (0.92–0.96)	0.87 (0.83–0.91)	0.90 (0.87–0.94)	0.85 (0.83–0.87)	0.89 (0.87–0.91)

[a]Raw alpha.
[c]Estimated on total score.
NODS = The NORC Diagnostic Screen for Gambling Problems (Gerstein et al., 1999), 30 days version.
PGSI = The Problem Gambling Severity Index (Ferris & Wynne, 2001).
PPGM = The Problem and Pathological Gambling Measure (PPGM; Williams & Volberg, 2013).

Table 6. Rasch analysis of PGSI, PPGM and NODS (N = 598).

	Item difficulty (SE)		Infit MNSQ		Outfit MNSQ	
	Dichotomized model[a]	Rating scale model	Dichotomized model[a]	Rating scale model	Dichotomized model[a]	Rating scale model
PGSI 1	0.05 (0.13)	1.65 (0.07)	0.77	0.85	0.53	0.73
PGSI 2	0.87 (0.14)	2.58 (0.07)	1.09	1.21	1.02	1.16
PGSI 3	−0.17 (0.13)	1.59 (0.07)	0.86	0.89	0.70	0.84
PGSI 4	1.45 (0.14)	2.72 (0.07)	0.82	0.98	0.67	0.78
PGSI 5	−0.13 (0.13)	1.43 (0.07)	0.64	0.74	**0.38**	0.58
PGSI 6	−0.02 (0.13)	1.65 (0.07)	0.80	0.83	0.59	0.74
PGSI 7	1.20 (0.14)	2.75 (0.08)	1.26	**1.55**	1.38	**1.66**
PGSI 8	0.76 (0.14)	1.95 (0.07)	0.72	0.88	0.54	0.60
PGSI 9	−0.33 (0.13)	1.30 (0.07)	0.85	0.95	0.64	0.83
Underfit		**>1.50**	0	1	0	1
Overfit		**<0.50**	0	0	1	0
EAP reliability	0.87	0.88				
Person reliability (*n* strata)	0.82 (2–3 strata)	0.82 (2–3 strata)				

Dichotomous Rasch model	Item difficulty (SE)	Infit MNSQ	Outfit MNSQ
PPGM 1	1.4 (0.13)	0.76	0.57
PPGM 2	1.12 (0.13)	0.71	0.50
PPGM 3	2.46 (0.13)	0.92	0.58
PPGM 4	3.08 (0.14)	0.89	0.58
PPGM 5	3.68 (0.15)	1.16	**1.82**
PPGM 6	5.11 (0.19)	1.09	**4.11**
PPGM 7	2.21 (0.13)	0.98	0.75
PPGM 8	0.80 (0.13)	0.65	**0.43**
PPGM 9	0.47 (0.13)	0.80	0.65
PPGM 10	2.51 (0.13)	1.44	1.42
PPGM 11	1.86 (0.13)	1.00	0.88
PPGM 12	0.52 (0.13)	0.89	0.74
PPGM 13	1.44 (0.13)	0.91	0.69
PPGM 14	2.33 (0.13)	1.08	1.09
EAP reliability	0.87		
Underfit	**>1.50**	0	2
Overfit	**<0.50**	0	1
Person reliability (*n* strata)	0.78 (2 strata)		

Dichotomous Rasch model[b]	Item difficulty (SE)	Infit MNSQ	Outfit MNSQ
NODS 1 or 2	1.50 (0.13)	0.90	0.79
NODS 3	3.34 (0.15)	1.01	0.86
NODS 5	2.41 (0.14)	0.95	0.82
NODS 7	2.93 (0.14)	0.92	0.63
NODS 8 or 9	2.36 (0.14)	0.91	0.77
NODS 10	1.90 (0.13)	0.78	0.64
NODS 12	3.50 (0.15)	0.84	**0.41**
NODS 13	4.61 (0.19)	0.91	0.50
NODS 14, 15 or 16	3.27 (0.15)	1.12	1.14
NODS 17	2.95 (0.14)	1.17	1.02
EAP reliability	0.75		
Underfit	**>1.50**	0	0
Overfit	**<0.50**	0	1
Person reliability (*n* strata)	0.50 (1–2 strata)		

[a]= The dichotomous model for the PGSI compared response alternative 0 (never) in relation to the rest, 1, 2 and 3 (Sometimes, Most of the time, and Almost always, collapsed).
[b]= Analysis was performed on items relevant for NODS scoring, of which some are collapsed (max item score = 1).
SE = Standard Error.
MNSQ = Mean-square.
EAP = Expected a posteriori.
NODS = The NORC Diagnostic Screen for Gambling Problems (Gerstein et al., 1999), 30 days version.
PGSI = The Problem Gambling Severity Index (Ferris & Wynne, 2001).
PPGM = The Problem and Pathological Gambling Measure (PPGM; Williams & Volberg, 2013).

Rasch analysis

Rasch analysis was used to estimate item difficulty, reliability and fit measures. Dichotomous models were used for the PPGM, the NODS and the PGSI. Due to use of multiple response alternatives in the PGSI, an additional rating-scale model was estimated which showed fewer items outside outfit thresholds in comparison with the dichotomous model (see Table 6). The PGSI had the smallest item difficulty range (1.30 to 2.85), where item 9 (guilt about gambling) was the smallest and item 2 (increased bets to get same excitement) the highest. NODS item difficulty ranged between 1.50 and 4.61, where item 1 or 2 (preoccupation) was the lowest and item 13 (gambling-related stolen money) highest. The PPGM had the largest item difficulty range (0.47 to 5.11), with $PPGM_{item9}$ (chasing losses) being the lowest and $PPGM_{item6}$ (gambling-related illegal acts) the highest. All measures had items showing under- or overfit, which might be problematic for measurement. Finally, person reliability indicated that PGSI and PPGM (0.82 and 0.78, respectively) could discriminate the sample into two or three levels, while NODS (0.50) could discriminate the sample into one or two levels. See Table 6.

Discussion

The aim of the current study was to test the factor structure and Rasch estimates of the PGSI, the PPGM, and the NODS among gamblers from recreational and help-seeking samples (see Table 7 for an overview of the results).

Table 7. Overview.

Factor structure	PGSI	PPGM	NODS
Model fit	Poor fit for the theoretically assumed single factor CFA model. EFA indicated a three factor solution.	Good fit for the theoretically assumed three factor CFA model, as well as the single factor CFA model.	Good fit for the theoretically assumed single factor CFA model.
Items <0.40	-	$PPGM_{item6}$ (both models)	-
Invariance	Excellent Cronbach's alpha across gender, age and gambling groups.	Excellent Cronbach's alpha across gender, age and gambling groups.	Good to excellent Cronbach's alpha across gender, age and gambling groups.
Rasch analysis			
Items showing infit or outfit	$PGSI_{item5}$, $PGSI_{item7}$	$PPGM_{item5}$, $PPGM_{item6}$, $PPGM_{item8}$	$NODS_{item12}$
Item difficulty range	1.30 to 2.85 (smallest)	0.47 to 5.11 (largest)	1.50 to 4.61
Person reliability (*n* gambling strata)	0.82 (2–3 strata)	0.78 (2 strata)	0.50 (1–2 strata)
Gambling severity levels in scoring index (*n*)	No problems, low at-risk, moderate at-risk, problem gambling ($n = 4$)	Recreational, at-risk, problem gambling, pathological gambling ($n = 4$)	No problems, at-risk, problem gambling, pathological gambling ($n = 4$)

PPGM = The Problem and Pathological Gambling Measure (PPGM; Williams & Volberg, 2013).
NODS = The NORC Diagnostic Screen for Gambling Problems (Gerstein et al., 1999), 30 days version.
PGSI = The Problem Gambling Severity Index (Ferris & Wynne, 2001).

Psychometric evaluation

CFA were used to test theoretically assumed constructs. For the assumed single-factor CFA model the PGSI performed worse and showed poor fit (RMSEA = 0.135, CFI = 0.948, TLI = 0.931, DF = 27, $\chi 2 = 320.1$; Bowen & Guo, 2011), compared to the PPGM (RMSEA = 0.042, CFI = 1.00, TLI = 0.96, DF = 77, $\chi 2 = 158.8$) and the NODS (RMSEA = 0.000, CFI = 1.00, TLI = 1.00, DF = 35, $\chi 2 = 28.5$). In contrast to how the PGSI is scored, that is, a total score, a follow-up exploratory factor analysis indicated that a three-factor solution of the PGSI showed best fit for the data. As previously reported, research of the PGSI factor structure has shown mixed results, mainly indicating unidimensionality, but also possibility of multiple factor structures (Holtgraves, 2009; Miller et al., 2013; Orford et al., 2010). The NODS showed good fit for the theoretically assumed unidimensional factor CFA model (RMSEA = 0.000, CFI = 1.00, TLI = 1.00, DF = 35, $\chi 2 = 28.5$). This finding is in line with previous studies of the NODS using explorative factor analysis (i.e. principal component analysis) (Hodgins, 2004; Toce-Gerstein et al., 2003). Given that the NODS is based on the previous DSM-IV diagnostic criteria our results support the notion of an unidimensional diagnosis of pathological gambling. The PPGM showed good fit for both the theoretically assumed three-factor CFA model (RMSEA = 0.030, CFI = 1.00, TLI = 1.00, DF = 74, $\chi 2 = 114.7$), and a unidimensional CFA model (RMSEA = 0.042, CFI = 1.00, TLI = 0.96, DF = 77, $\chi 2 = 158.8$). To the best of our knowledge (and personal communication Rachel Volberg, 1 July 2021) the factor structure of the PPGM has previously not been evaluated. As such, these findings remain to be corroborated in future studies. In the current study, the only individual CFA item factor loading <0.40 (i.e., a commonly used reference threshold for low factor loading), was $\mathrm{PPGM}_{\mathrm{item6}}$ which assesses illegal acts to finance gambling. Furthermore, $\mathrm{PPGM}_{\mathrm{item6}}$ and $\mathrm{NODS}_{\mathrm{item13}}$ (which assesses taking money) had the highest item difficulties in the Rasch analysis. These findings are in line with previous research showing that removal of the illegal acts criterion increased the amount of explained variance in the factor structure, as well as classification accuracy (Petry et al., 2013). As previously mentioned, this criterion was removed in the DSM5 diagnostic criteria (American Psychiatric Association, 2013).

The PGSI, the PPGM, and the NODS are all used to assess problem gambling. The PGSI scoring includes two additional at-risk gambling levels and the PPGM and the NODS scoring for pathological gambling according to DSM-IV. In the Rasch analysis, Person reliability showed that no measure had sufficient number of strata in relation to the suggested severity levels in the scoring indexes, but the PGSI performed better than the other measures in this sense. The Rasch analysis also showed that all measures had items showing infit or outfit. Finally, the Rasch analysis showed that the PPGM had the largest item difficulty range and the PGSI the smallest. The PPGM item difficulty range seemed sufficient to detect the suggested gambling levels in the scoring index. For the NODS, the difficulty range seemed adequate to detect problems in the higher end of the spectra. This indicated support for using the NODS to identify pathological gambling in clinical samples, but also potential issues for detecting at-risk gambling. The narrow difficulty range for the PGSI was problematic in relation to the measure’s ability to identify low severity problems. Hence, it is unclear whether the PGSI can be used to detect such sub-clinical risk levels. Similar conclusions regarding the PGSI have also been reported in previous studies (Cowlishaw et al., 2019; Miller et al., 2013).

Strengths and limitations

The current study might be the first to test the factor structure of the PPGM. The study expanded analysis methods for gambling measures beyond classical test theory, and used item response theory estimates such as item difficulty range or person reliability. Further strengths included evaluation of individual items within the respective measures and not by adding items into a pool and evaluating problem gambling as a construct on an aggregated level. Finally, evaluations were performed within the same sample, enabling contrasting of results between measures. Such comparisons are scarce, particularly among European gambling samples. Concerning limitations, the sample in the study was relatively small compared to other gambling Rasch studies (Cowlishaw et al., 2019; Miller et al., 2013; Molde et al., 2010). Also, several participants had missing data on the NODS due to an administrative error. Another limitation is the focus on a relatively narrow psychometric area, leaving other questions for future studies (e.g. validation of cutoffs/diagnostic accuracy, predictive validity, and measurement invariance). Future studies should focus on further improving the PPGM with respect to the weak items presented in this study. The novel DSM-5 criteria call for new cutoffs and assessment of severity levels in gambling measures which stands as a suggestion for future research.

Conclusions and clinical implications

None of the measures performed perfect on all tests. The drawbacks of the PGSI were that it was unclear whether it measures one gambling construct or several, and that the instrument failed to discriminate mild subclinical gambling levels, such as at-risk gambling. Overall, the PPGM performed best of the three instruments. We conclude that the PPGM can be used in general populations and clinical contexts to detect problem gambling and pathological gambling, across a severity continuum. However, several PPGM items were weak in the CFA and Rasch analysis with respect to their psychometric properties. Finally, we conclude that the NODS is suitable for use in clinical samples for identification of pathological gambling. However, the Rasch analysis indicated that the NODS might be weak in discriminating severity levels, in particular subclinical pathological gambling problems.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The work was supported by the Development funds for identification and treatment of problem gambling from the Stockholm Health Care Services, Stockholm Region [NA]; The Swedish program grant “Responding to and Reducing Gambling Problems—Studies in Help-seeking, Measurement, Comorbidity and Policy Impacts” (REGAPS), financed by Forte, the Swedish Research Council for Health, Working Life and Welfare [2016-0709].

ORCID

Olof Molander http://orcid.org/0000-0001-5348-051X
Peter Wennberg http://orcid.org/0000-0002-0013-2965

Ethical approval

The regional ethical board of Stockholm 2017/1479-31.

Supplementary

Supplemental data can be accessed by online.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.).

Blaszczynski, A., & Nower, L. (2002). A pathways model of problem and pathological gambling. *Addiction*, *97*(5), 487–499. https://doi.org/10.1046/j.1360-0443.2002.00015.x

Bond, T. G., & Fox, C. M. (2013). *Applying the Rasch model: Fundamental measurement in the human sciences*. Psychology Press.

Bowen, N. K., & Guo, S. (2011). *Structural equation modeling*. Oxford University Press.

Calado, F., & Griffiths, M. D. (2016). Problem gambling worldwide: An update and systematic review of empirical research (2000-2015). *Journal of Behavioral Addictions*, *5*(4), 592–613. (PubMed). https://doi.org/10.1556/2006.5.2016.073

Cappelleri, J. C., Lundy, J. J., & Hays, R. D. (2014). Overview of classical test theory and item response theory for quantitative assessment of items in developing patient-reported outcome measures. *Clinical Therapeutics*, *36*(5), 648–662. https://doi.org/10.1016/j.clinthera.2014.04.006

Christensen, D. R., Williams, R. J., & Ofori Dei, S. M. (2019). The multidimensional structure of problem gambling: An evaluation of four gambling categorization instruments from an international online survey of gamblers. *Journal of Gambling Studies*, *35*(4), 1079–1108.

Cowlishaw, S., Merkouris, S. S., Dowling, N. A., Rodda, S., Suomi, A., & Thomas, S. L. (2019). Locating gambling problems across a continuum of severity: Rasch analysis of the quinte longitudinal study (QLS). *Addictive Behaviors*, *92*, 32–37. https://doi.org/10.1016/j.addbeh.2018.12.016

Enders, C. K., Mistler, S. A., & Keller, B. T. (2016). Multilevel multiple imputation: A review and evaluation of joint modeling and chained equations imputation. *Psychological Methods*, *21*(2), 222. https://doi.org/10.1037/met0000063

Ferris, J., & Wynne, H. (2001). The Canadian problem gambling index: Final report.

Gerstein, D., Volberg, R. A., Toce, M., Harwood, H., Johnson, R., Buie, T., Christiansen, E., Chuchro, L., Cummings, W., & Engelman, L. (1999). *Gambling impact and behavior study: Report to the national gambling impact study commission*. National Opinion Research Center.

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. https://doi.org/10.1016/j.addbeh.2004.03.017

Holtgraves, T. (2009). Evaluating the problem gambling severity index. *Journal of Gambling Studies*, *25*(1), 105. https://doi.org/10.1007/s10899-008-9107-7

Kang, H. (2013). The prevention and handling of the missing data. *Korean Journal of Anesthesiology*, *64*(5), 402–406. https://doi.org/10.4097/kjae.2013.64.5.402

Linacre, J. M. (2002). What do infit and outfit, mean-square and standardized mean?. *Rasch Measurement Transactions*, *16*(2), 878.

Merkouris, S. S., Greenwood, C., Manning, V., Oakes, J., Rodda, S., Lubman, D., & Dowling, N. A. (2020). Enhancing the utility of the problem gambling severity index in clinical settings:

Identifying refined categories within the problem gambling category. *Addictive Behaviors*, *103*, 106257. https://doi.org/10.1016/j.addbeh.2019.106257

Miller, N. V., Currie, S. R., Hodgins, D. C., & Casey, D. (2013). Validation of the problem gambling severity index using confirmatory factor analysis and rasch modelling. *International Journal of Methods in Psychiatric Research*, *22*(3), 245–255. https://doi.org/10.1002/mpr.1392

Molander, O., Volberg, R., Sundqvist, K., Wennberg, P., Månsson, V., & Berman, A. H. (2019). Development of the gambling disorder identification test (G-DIT): Protocol for a delphi method study. *JMIR Research Protocols*, *8*(1), e12006. https://doi.org/10.2196/12006

Molander, O., Wennberg, P., & Berman, A. H. (2021). The gambling disorders identification test (GDIT): Psychometric evaluation of a new comprehensive measure for gambling disorder and problem gambling. *Assessment*, 10731911211046044. https://doi.org/10.1177/10731911211046045

Molde, H., Hystad, S. W., Pallesen, S., Myrseth, H., & Lund, I. (2010). Evaluating lifetime NODS using Rasch modelling. *International Gambling Studies*, *10*(2), 189–202.

Orford, J., Wardle, H., Griffiths, M., Sproston, K., & Erens, B. (2010). PGSI and DSM-IV in the 2007 British gambling prevalence survey: Reliability, item response, factor structure and inter-scale agreement. *International Gambling Studies*, *10*(1), 31–44. https://doi.org/10.1080/14459790903567132

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. https://doi.org/10.1111/j.1360-0443.2012.04087.x

R Core Team. (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Toce-Gerstein, M., Gerstein, D. R., & Volberg, R. A. (2003). A hierarchy of gambling disorders in the community. *Addiction*, *98*(12), 1661–1672. https://doi.org/10.1111/j.1360-0443.2003.00545.x

Wickwire, E. M., Burke, R. S., Brown, S. A., Parker, J. D., & May, R. K. (2008). Psychometric evaluation of the national opinion research center DSM-IV screen for gambling problems (NODS). *American Journal on Addictions*, *17*(5), 392–395. https://doi.org/10.1080/10550490802268934

Williams, R. J., & Volberg, R. A. (2013). The classification accuracy of four problem gambling assessment instruments in population research. *International Gambling Studies*, *14*(1), 1–14. https://doi.org/10.1080/14459795.2013.839731

Williams, R. J., Volberg, R. A., & Stevens, R. M. G. (2012). *The population prevalence of problem gambling: Methodological influences, standardized rates, jurisdictional differences, and worldwide trends* [Technical Report]. Ontario Problem Gambling Research Centre. https://opus.uleth.ca/handle/10133/3068

Wilson, M. (2004). *Constructing measures: An item response modeling approach: An item response modeling approach*. Routledge. https://doi.org/10.4324/9781410611697

Wind, S., & Hua, C. (2021). Rasch measurement theory analysis in R: Illustrations and practical guidance for researchers and practitioners. Bookdown. org,[Epub].

Wright, B. D., & Linacre, J. M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions*, *8*(3), 370.