International Journal of Methods in Psychiatric Research
Int. J. Methods Psychiatr. Res. 22(3): 245–255 (2013)
Published online 9 September 2013 in Wiley Online Library
(wileyonlinelibrary.com) **DOI:** 10.1002/mpr.1392

Validation of the problem gambling severity index using confirmatory factor analysis and rasch modelling

NATALIE V. MILLER,[1] SHAWN R. CURRIE,[2] DAVID C. HODGINS[2] & DAVID CASEY[2]

1 University of British Columbia, Vancouver, BC, Canada
2 University of Calgary, Calgary, AB, Canada

Key words
problem gambling, PGSI, assessment, factor analysis, Rasch modelling

Correspondence
Shawn R. Currie, Mental Health Information & Evaluation Unit, Alberta Health Services, Addiction and Mental Health Services, 10101 Southport Rd SW, Calgary, AB T2W 3N2, Canada.
Telephone (+1) 403 703-4931
Fax (+1) (403) 943-0199
Email: scurrie@ucalgary.ca

Received 9 January 2012; revised 5 June 2012; accepted 19 June 2012

Abstract

The Problem Gambling Severity Index (PGSI), a screening tool used to measure the severity of gambling problems in general population research, was subjected to confirmatory factor analysis and Rasch modelling to (a) confirm the one-factor structure; (b) assess how well the items measure the continuum of problem gambling severity; (c) identify sources of differential item functioning among relevant subpopulations of gamblers. Analyses were conducted on a nationally representative sample of over 25,000 gamblers compiled by merging data from the Canadian Community Health Survey and Canadian Problem Gambling Index (CPGI) integrated datasets. Results provided support for a one-factor model that was invariant across gender, age, income level, and gambler type. Rasch modelling revealed a well-fitting, unidimensional model with no miss-fitting items. The average severity assessed by the PGSI is consistent with moderately severe problem gambling. The PGSI is therefore weak in assessing low to moderate problem severity, a notable limitation of most brief gambling screens. Evidence of clinically significant differential item functioning was found with only one item, borrowing money to gamble, which behaved differently in gamblers who play electronic gaming machines or casino games compared to gamblers who avoid these games. *Copyright © 2013 John Wiley & Sons, Ltd.*

Introduction

Given the rapid expansion of legalized gambling in North America over the past 15 years, government policy-makers require regular population surveillance of problem gambling and associated harms. Several brief screening tools are available to monitor rates of problem gambling although research on their psychometric properties is often limited. Among the available tools, the Southern Oaks Gambling Screen (SOGS) (Lesieur and Blume, 1987), the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) based measures of problem gambling, such as the National Opinion Research Center DSM-IV Screen for Gambling Problems (NODS) (Gerstein *et al.*, 1999), and the Problem Gambling Severity Index (PGSI) of the Canadian Problem Gambling Index (CPGI) (Ferris and

Wynne, 2001) are frequently used for population-level measurement of problem gambling.

The SOGS, a 20-item measure, was initially developed as a tool for clinical assessment and later applied to survey research, but without extensive validation using a community sample. Consequently, the SOGS has been criticized for its ability to accurately assess problem gambling within the general population (Stinchfield *et al.*, 2007). The SOGS tends to over-estimate problem gambling within community samples compared to other measures including the PGSI (Ferris and Wynne, 2001; Stinchfield, 2002). Indeed, an item response analysis of the SOGS found that community samples responded differently to several SOGS items compared to pathological gamblers (Strong *et al.*, 2003). The NODS, a 34-item measure, was constructed to mirror DSM-IV criteria for pathological gambling and therefore, like the SOGS, the content of this scale reflects the clinical range of pathological gambling. Consequently, the NODS has low sensitivity for detecting mild levels of problem gambling severity (Abbott and Volberg, 2006; Orford *et al.*, 2010).

As screening tools for pathological gambling, the SOGS and NODS have shown to possess robust properties (Stinchfield *et al.*, 2007) but may be limited in characterizing milder or subthreshold forms of problem gambling in the general population. The PGSI was specifically designed and validated as a quantitative index of problem gambling severity within the general population (Ferris and Wynne, 2001). The PGSI also was created to identify gamblers at risk for developing problem gambling. Hence, the PGSI assesses "soft" indicators of problem gambling rather than focusing on diagnostic criteria. Items assess both gambling behaviour and the financial, social, and health consequences of gambling. The PGSI has strong internal consistency, test–retest reliability, concurrent validity with other problem gambling measures, and criterion validity with measures of gambling involvement (Ferris and Wynne, 2001; Holtgraves, 2009; Mcmillen and Wenzel, 2006; Neal *et al.*, 2005). An advantage of the PGSI is the four-point ordinal scale for each item rather than the dichotomous (yes/no) format used with other measures. The total score and corresponding interpretative categories [non-problem gamblers (0), low-risk (1–2), moderate-risk (3–7), problem-gambler (> 7)] provide full coverage of the continuum of problem gambling severity. The non-problem, low-risk, and moderate-risk gambler categories represent a meaningful progression in level of risk for all gamblers who are below the threshold of problem gambling, a property of the scoring system we validated in a recent population study by our research group (Currie *et al.*, 2013). The combination of item diversity and ordinal scaling allows the PGSI to tap levels of problem gambling that may be overlooked by the dichotomous-scored, diagnostically-focused measures of pathological gambling.

Since its development in 2001, the PGSI has become the gold standard in both Canada and Australia for collecting information on problem gambling symptoms (McCready and Adlaf, 2006; Neal *et al.*, 2005), and has been used worldwide in population-level survey research [Department of Social Development (Northern Ireland), 2010; Kavli and Berntsen, 2005; Kristiansen and Jensen, 2011; Ministry of Health, 2009; Olason and Gretarsson, 2009; Volberg and Bernhard, 2006; Wardle *et al.*, 2007; Young and Stevens, 2008]. Given both the strengths and gaining popularity of the PGSI, it is important to empirically validate all dimensions of its psychometric properties. Specifically, there is a gap in the literature between theoretical features of the PGSI, as a measure assessing the full continuum of the single construct problem gambling, and empirically supported aspects of the measure, such as its factor structure and how well the individual items represent the problem gambling continuum. Although a handful of studies have investigated the factor structure of the PGSI, no study has submitted the PGSI to a comprehensive item response analysis (Orford *et al.*, 2010).

To date, factor analytic studies of the PGSI have produced a single-factor solution (Holtgraves, 2009; Orford *et al.*, 2010). Using data from the 2007 British Gambling Prevalence Survey and principal components analysis (PCA), Orford *et al.* (2010) determined a one-factor model (accounting for 62% of the variance) was the best fit for the PGSI. A PCA on Canadian population data also produced a one-factor structure accounting for 49% of the variance (Holtgraves, 2009). The factor analytic approach in these studies relied exclusively on exploratory methods, which are more appropriate for developing measures (Hoyle and Smith, 1994). Confirmatory factor analysis (CFA), involving the fitting of a pre-determined structure to a dataset, is a more restrictive test of factorial validity than exploratory factor analysis (Currie *et al.*, 2004; Hoyle and Smith, 1994). Thus, the first aim of the current study was to assess the factor structure of the PGSI using CFA methods.

Previous research supports the PGSI's coverage of the problem gambling continuum albeit in a limited fashion. Orford *et al.* (2010) examined the frequency of item endorsement across PGSI problem gambling categories, and concluded the items adequately captured the problem gambling continuum. Holtgraves (2009), using similar methodology, reported certain items were more characteristic of low-risk gamblers (i.e., chasing losses, feeling guilty, betting more than could afford) than other items (i.e., borrowing money or financial problems due to gambling). These results suggest that individual PGSI items assess different levels of problem gambling. However, these studies relied upon descriptive statistics and visual inspection of the data to support their conclusions. The present study aimed to improve on these subjective methods by using item response analysis.

Rasch modelling is one form of item response analysis that assesses the dimensionality of questionnaire measures, providing an evaluation of the fit between items and the underlying latent construct (Strong *et al.*, 2004b). This model analyses item response patterns to characterize the performance of individual items in measuring the underlying construct of problem gambling. There are three notable features of this model. First, people and items are independently measured along the same latent construct (i.e., problem gambling continuum) using the metric of an equal interval logit scale (Wright and Stone, 1979). The position of each item along the latent problem gambling continuum is defined by a severity parameter. This parameter indicates the point along the continuum where there is a 50% chance of item endorsement (Embretson and Reise, 2000).

Second, the presence of more severe gambling problems presumes the presence of less severe problems in the same individual. By default, zero on the latent problem gambling continuum is set as the mean level of problem gambling severity across all items. Severity estimates greater than zero reflect items that assess problem gambling that are more severe than the mean level of severity for the PGSI scale. The rank order of symptoms by severity estimate provides an indication of the specific types of problems that are more likely to appear as problem gambling severity increases.

Finally, infit and outfit χ^2 statistics provide information about how well items fit within the model (e.g., high severity items endorsed by low severity individuals are poor fits; Linacre, 2003). The infit statistic provides information about inlier-sensitive fit, that is, the degree of fit between items and people (i.e., the model predicts that low severity gambler endorse low severity items). Whereas the outfit statistic describes the degree of outlier-sensitive fit (i.e., the degree to which there is a mismatch between items and people in terms of severity). A χ^2 above 1.0 indicates the model is unpredictable or noisy; values less than 1.0 signal that the data is too predictable or redundant (Linacre, 2003).

In addition to these elements, another feature of item response analysis is the ability to examine differential item functioning (DIF), which occurs when subgroups respond uniquely to items (Molde *et al.*, 2010). For example, younger age was associated with reporting "chasing losses" at a lower level of problem gambling severity compared to older age in an item analysis of the NODS (Strong and Kahler, 2007). An assumption of all measures of problem gambling is that certain scores reflect certain degrees of problem severity across all individuals. When DIF occurs this assumption is violated and consequently affects the psychometric validity of the measure. Although other problem gambling measures have been evaluated using Rasch modelling and DIF analyses, this analysis has not been applied to the PGSI (Strong *et al.*, 2003; Strong and Kahler, 2007).

The application of DIF procedures is relevant to the PGSI in light of the variability in problem gambling prevalence rates across demographic and gambling type subgroups. Individuals who are male, younger, or earn lower incomes tend to endorse more PGSI items compared to individuals who are female, older, or earn higher incomes (Cox *et al.*, 2005; Orford *et al.*, 2010). Furthermore, individuals who gamble at casinos or with electronic gaming machines also are more likely to report gambling problems compared to other types of gamblers (Currie *et al.*, 2008c). Presently, differences in problem gambling prevelance are interpreted as actual differences in problem gambling (Cox *et al.*, 2005). Nonetheless, it is possible that these differences are due to instrument bias. Clearly this issue must be disambiguated before theoretical arguments about gambling prevalence, such as potential gender differences, can be empirically supported by PGSI data. Therefore, another aim of the current study was to analyse PGSI items using DIF procedures. Because statistical tests should not alone determine the validity of an instrument, we also incorporated feedback from experts in problem gambling as a measure of the continuum of problem gambling severity.

Present study objectives

The specific objectives of this study were to (1) assess the factor structure of the PGSI using confirmatory methods with the largest sample of completed PGSI surveys available for research to date; (2) model the continuum of gambling problems as assessed by the PGSI total score with Rasch modelling; (3) identify sources of DIF across demographic and game type subgroups; (4) assess the construct validity of the PGSI's item coverage by comparing item severity estimates with problem gambling expert opinions of item rankings. For objective one, we predicted the one- factor model would be validated by CFA on the basis of strong fit metrics and high internal consistency. For objective two, we predicted low severity items and high severity items would be consistent with previous research (Holtgraves, 2009). But, as the PGSI has not been submitted to this analysis before, we did not predict a specific pattern of severity across items. Similarly, we had no a priori predictions for objective three. Regarding objective four, we anticipated that expert rankings would be congruent with the severity estimates generated in the rasch model.

Method

Data sources

We used data on the PGSI derived from two large population datasets, which we merged for the purposes of this study: the CPGI integrated dataset and the Canadian Community Health Survey. The full CPGI collects data on both problem gambling symptoms and gambling habits (frequency, time spent, and expenditures in Canadian dollars) for all major types of gambling (e.g. lottery, casino games, bingo, internet, electronic gaming machines) (Ferris and Wynne, 2001).

CPGI integrated dataset

The CPGI integrated dataset includes data from prevalence surveys conducted in six provincial and one national survey for a combined sample size of 14,388 past year gamblers. We have used this dataset in previous research (Currie *et al.*, 2008a; Currie *et al.*, 2006; Currie *et al.*, 2008b; Currie *et al.*, 2013; Miller and Currie, 2008). Each survey employed a similar methodology with the data collected via a computer-assisted telephone interviewing system. Samples were stratified by regional boundaries, age, and gender. Random digit dialling was used to select cases at the household level and the adult resident ($\geq$ 18 years) with the most recent birthday was selected to be interviewed. The prevalence of problem gambling (PGSI $>$ 7) ranged from 0.4% to 1.3% across surveys.

Canadian Community Health Survey – Mental Health Well-being (CCHS-1.2)

Detailed descriptions of the Canadian Community Health Survey – Mental Health Well-being (CCHS-1.2) in terms of target population, sampling procedures, response rate, and psychiatric assessment are in other sources (Currie *et al.*, 2006; Gravel and Beland, 2005; Statistics Canada, 2002). Briefly, the CCHS-1.2 was a 2002 cross-sectional survey of a nationally representative sample of individuals aged 15 and older. Participants were recruited using the same basic random digit dialling method as in the CPGI integrated dataset. Valid PGSI data for past year adult gamblers were available for 18,913 cases.

Both datasets were restricted to individuals who were (a) $\geq$ 18 years of age, (b) reported gambling at least once in the past year, and (c) administered the PGSI items. The final merged dataset included demographic, gambling involvement, and PGSI data for 25,584 cases. To test factorial invariance and DIF by game preference, we collapsed the CPGI data on games played into two categories: (1) gaming machine or casino gamblers were persons who reported playing electronic gaming machines [EGMs; video lottery terminals (VLTs) and slot machines] or casino games in the past 12 months; (2) other gamblers consisting of individuals who reported types of gambling other than EGMs or casino games (e.g., lottery tickets, raffle tickets). Research has shown that EGM and casino gambling have a greater risk of harm and addiction compared to other forms of gambling (Breen and Zimmerman, 2002; Welte *et al.*, 2004) and we have used this distinction in prior studies (Currie *et al.*, 2006; Currie *et al.*, 2008b; Miller and Currie, 2008; Currie *et al.*, 2013). Table 1 compares the demographics of the two CPGI datasets. Participants in the CPGI integrated and CCHS-1.2 were largely comparable in demographic and gambling profile. The CCHS-1.2 contained a higher proportion of males ($\chi^2 = 10.43$, $p < 0.01$), older adults (65+ years; $\chi^2 = 233.64$, $p < 0.001$), and individuals earning less than $30,000CAN annual income ($\chi^2 = 283.24$, $p < 0.001$). These differences, although significant because of the very large sample sizes, were not large in magnitude. There were no differences in education, proportion of EGM/casino players, or mean PGSI scores between the groups (all p values > 0.05).

Survey of gambling experts

A web-based survey on assessing problem gambling severity with the PGSI was completed by 142 gambling experts (researchers, clinicians, and government and industry persons) who were based in Canada (78%), United States (14%), and other countries (8%). Respondents were recruited from several sources: a mailing list used in a prior gambling expert survey (Currie *et al.*, 2008a), lists of attendees at major gambling conferences in the past five years, researchers who received funding for gambling from Canadian and American granting agencies, and members of the Gambling Issues International Listserv. An email invitation was sent to 422 individuals making the response rate 34%. A web-based survey was developed and posted online. Gambling experts were asked to rate the severity of each PGSI item level in terms of providing information on overall severity of problem gambling being exhibited by gamblers who endorse the symptom. Specifically, experts rated whether endorsement of the item (as occurring sometimes, most of the time, or almost always) was indicative of the gambler having no gambling problem (1), a mild problem (2), a moderate problem (3), or a severe gambling problem (4). The final ranking was derived by

Table 1. Demographics of two datasets used in analyses

Characteristic	CPGI integrated (Currie *et al.*, 2008b)	CCHS-1.2 (Currie *et al.*, 2006)
N	14388	18913
Age (%)		
18–24 years	9.6	11.2
25–64 years	78.0	69.9
65+ years	12.4	18.9
Gender (% male)	48.7	50.7
Played EGM or casinos in past year (%)	44.5	44.1
Income		
< \$30 K	20.5	30.0
\$30 K–\$50 K	24.7	24.1
> \$50 K	54.8	45.9
Mean PGSI score (standard deviation)	0.5 (1.8)	0.5 (1.9)

Note: All datasets excluded non-gamblers and persons under 18 years.

the proportion of experts who endorsed the item as a moderate or severe problem gambling symptom. Additionally, the average of experts' severity ratings was calculated for each item.

Data analysis

Factor analysis

CFA was conducted using AMOS V17. Standard fit indices included the comparative fit index (CFI with values greater than 0.95 indicating a good fit) and the root mean square error of approximation (RMSEA; values less than 0.06 indicate a good fitting model) (Hu and Bentler, 1999). Internal consistency of the summed total score was estimated using coefficient alpha. Factor invariance across gender, age, type of game, and income was assessed using CFA of the total data file.

Rasch modelling

A rating scale Rasch model was used to: (1) estimate the relative severity of problem gambling being assessed by each item; (2) assess the order in which symptoms covered by the PGSI are likely to be endorsed as severity increases; (3) assess the standard error of measurement associated with each item and the full range of PGSI scores. The analysis was conducted on the original four-point scaled PGSI items. With this approach, the Rasch-generated severity estimates represent the overall severity for the item across all non-zero responses (Bond and Fox, 2007). Rasch modelling requires variability in responses across items therefore individuals who endorsed no PGSI items (n = 21,597) or all PGSI items (n = 5) were removed from the analysis, leaving 3992 cases.

The PGSI was evaluated using a Rasch model in WINSTEPS. WINSTEPS also was used to assess PGSI items for DIF by comparing individual item severity estimates between subgroups (Wright and Stone, 1979). In addition to statistically significant differences identified through WINSTEPS, we also considered the clinical significance of differences by examining the magnitude of severity estimate differences between subgroups. Significant DIF findings were identified on the basis of two criteria: (1) the result of the *t*-test comparing severity estimates was significant at the $p < 0.01$ level, and (2) the difference in estimates was 0.5 logit units or greater, which represents a half standard deviation difference (Linacre, 2003; Steinberg and Thissen, 2006).

Results

Confirmatory factor analysis (CFA)

A one-factor structural equation model of the PGSI Likert response category items showed good fit to the data, despite a significant χ^2 (27) = 2393.6 {CFI = 0.972; RMSEA = 0.059 [90% confidence interval (CI) 0.057–0.061]}. Because item data were positively skewed, bootstrapping was used to estimate parameter CIs. Factor loadings ranged from 0.54 (borrowing money) to 0.77 (acknowledge problem with gambling). Corrected item-total correlations for the summed total score ranged from 0.49 to 0.70 for the PGSI items (M r = 0.61), with a coefficient alpha of 0.86. The coefficient alpha did not increase with the deletion of any item.

Factor invariance based on gender (males versus females), age (18–24 years, 25–64 years, and 65+ years), household income (< \$30 K, \$30 K–\$50 K, > \$50 K; all currency Canadian), and game type (EGM/casino versus other) was assessed using CFA of the complete sample. These analyses revealed factor invariance across all subgroups with only minor differences in item loadings. There was no evidence of meaningful differences between men and women, types of gamblers or, among levels of income. There were some statistical differences among the three age groups, with slightly lower loadings for older participants. Detailed descriptions of these analyses can be provided by the authors.

Rasch modelling

Testing the assumptions of unidimensionality and local independence

CFA results confirmed a single factor underlying the PGSI, supporting the Rasch model assumption of unidimensionality of the latent construct (Langenbucher *et al.*, 2004). After the Rasch calibration, a PCA of the residuals revealed no violations of local independence (i.e., responses to one item were independent of responses to other items). The eigenvalue of the first residual factor was 1.4 units. The plot of the standardized residuals exhibited a random pattern suggesting no additional factors. The largest correlation between the residuals of two items was 0.29, further indication of local independence.

Model fit

The rating scale Rasch model, conducted on the original four-point scaled items, was deemed appropriate on the basis of: (1) a smooth distribution of category frequencies across items (although all items were positively skewed, the distribution was consistent across items); (2) the average measures for each item category (i.e., the average of the severity estimates for all persons in the sample who chose that response category) showed advancement in value with each successive rating point, showing that higher problem gambling severity is associated with higher category labels (Smith *et al.*, 2003); (3) no indication of disordered items in terms of severity (i.e., the average measure of a higher score value is lower than the average measure of lower scale value; Linacre, 1999). In light of these observations, the comparison of each item's overall severity value is valid with polytomous data results.

All of the infit and outfit statistics for individual PGSI items fell within the acceptable range of 0.5 to 1.5 (Wright and Linacre, 1994) indicating a good model fit for all items. The standard errors fell within a narrow range of 0.03 to 0.05 for all PGSI items including those that were endorsed infrequently by gamblers (e.g. borrowing money). The reliability estimate for item severity indicated perfect reliability (1.00) suggesting that the item rankings for severity would likely replicate in other similar population samples (Wright and Stone, 1979).

Item-level estimates of problem severity

Table 2 presents the item characteristics for all nine PGSI items. The most commonly reported symptoms were "feeling guilty", "chasing losses", and "betting more than one could afford to lose". These items also were typical of low levels of problem gambling severity, with severity estimates clustering at $\delta = -0.75$ to -0.80. Items typical of high levels of severity included "health problems", "financial problems", and "borrowing money". In contrast to the clustering of low severity items, these items reflected a range of problem gambling severity ($\delta = 0.45$ to 1.17). Finally, moderate levels of problem gambling were represented by a cluster of items at $\delta = 0$ to 0.05. Thus, this analysis revealed a gap in PGSI item coverage between mild and moderate problem gambling, and identified that redundant severity information was provided by items within the clusters at mild and moderate problem gambling. Ideally, severity estimates should be at least 0.15 logit units apart (Garrett, 2003) to avoid redundant psychometric information being provided by two or more items. The items "feeling guilty" and "chasing losses" appear to provide redundant information on problem severity.

Differential item functioning (DIF)

Subgroup comparisons were identical to those used in the investigation of factor structure invariance: gender, age, household income, and game type. Complete DIF results are shown in Table 3. Although several items had a statistically significant difference in severity estimates, only "borrowing money" met both criteria for significance testing. Specifically, EGM or casino gamblers had a significantly higher severity estimate than other gamblers ($\delta = 1.34$ versus $\delta = 0.70$).

Severity rankings compared to expert opinion

We compared the empically-derived rank order of symptoms to the opinions of gambling experts collected from the online survey (Table 4). The intraclass correlation coefficient between these measures was 0.79, suggesting good agreement between rankings. The average severity rating by experts for each item was also highly correlated

Int. J. Methods Psychiatr. Res. 22(3): 245–255 (2013). DOI: 10.1002/mpr

Table 2. Symptom severity estimates for the Problem Gambling Severity Index (PGSI)[a] from gamblers who reported gambling at least once in the previous 12 month period

		Problem Gambling Severity		Infit statistic		Outfit statistic		
Item	Percentage endorsed[b]	δ	Standard error	M	Z	M	Z	r_t[c]
Guilt feelings	6.8	−0.85	0.03	0.99	−0.4	1.03	0.9	0.73
Chasing losses	7.0	−0.85	0.03	1.13	4.4	1.18	5.5	0.71
Betting more than can afford	6.0	−0.70	0.03	1.12	3.8	1.15	4.5	0.69
Recognizes one has a problem	3.6	0.00	0.03	0.92	−2.3	0.72	−7.7	0.65
Tolerance	3.9	0.00	0.03	1.20	5.3	1.12	3.0	0.59
Criticized by others	3.6	0.05	0.03	1.15	3.9	1.05	1.3	0.59
Health problems	2.7	0.48	0.04	1.07	1.7	0.93	−1.3	0.55
Financial problems	2.2	0.70	0.04	0.93	−1.7	0.58	−9.1	0.56
Borrowing money	1.7	1.17	0.05	1.16	3.0	0.99	−0.1	0.45

[a]The full PGSI can be found online at www.problemgambling.ca/EN/Documents/ProblemGamblingSeverityIndex.pdf.
[b]Proportion with non-zero score.
[c]r_t point-biserial correlation between item and total measured severity level based on Rasch calibrated scores.

Table 3. Results of DIF tests of statistical and clinical significance

	DIF			
PGSI item	Gender	Age	Income	Game type
Guilt feelings	F–M = −0.28			
Chasing losses	F–M = 0.20	18–24 years–overall = −0.24		Other–EGM/ casino = −0.28
Betting more than can afford	F–M = −0.21		> $50 K–overall = 0.20	
Recognizes one has a problem		18–24 years–overall = 0.44		Other–EGM/ casino = 0.34
Tolerance			$30 K–$50 K–overall = −0.15> $50 K–overall = 0.22	
Criticized by others				
Health Problems		18–24 years–overall = −0.10		
Financial Problems	F–M = 0.23			
Borrowing money				*Other–EGM/ casino = −0.65*

Note: Cells with problem gambling severity parameters reported indicate the difference between groups ($p < 0.01$). The difference between the subgroups is shown for cells with significant results. Gender and Game Type were two-group comparison. For Age and Income the entire sample was the reference group and the specified category was the focal group. Italic typeface items indicate a severity parameter difference of 0.50 or greater.

with the Rasch model item-level severity estimates, $r = 0.82$, $p < 0.01$. Agreement was best at the upper and lower extremes of problem gambling severity. Congruent with our empirical results, experts reported that “borrowing money” and “financial problems” were most indicative of severe gambling problems, whereas “feeling guilty” was least indicative of severe gambling problems. Furthermore, “health problems” was found to be a mid-level symptom

Table 4. Severity rankings of the nine PGSI Items: Results from Rasch analysis and expert survey

PGSI symptom	Average severity rating by experts[a]	Percentage of experts rating symptom as severe or moderate	Severity rank: Rasch model	Severity rank: Experts
Borrowing money	3.43	83	*1*	*1*
Financial problems	3.44	82	*2*	*2*
Health problems	3.23	75	*3*	*4*
Criticized by others	3.11	70	4	8
Tolerance	3.10	71	*5*	*7*
Recognizes one has a problem	3.19	73	*6*	*6*
Betting more than can afford	3.13	74	7	3
Chasing losses	3.17	71	8	5
Feelings of guilt	2.93	60	*9*	*9*

[a]Scale: 1 = no gambling problem; 2 = mild problem; 3 = moderate problem; 4 = severe problem.
Note: Rasch model and expert rankings that agree within two scale points shown in italic typeface.

of problem gambling severity by both methods. In contrast, there was a large disagreement between experts and the Rasch model for “chasing losses”. According to the model, this item is typical of low levels of severity, but experts rated this item as being indicative of more severe problem gambling. A similar disagreement occurred with “criticized by others”. In this case experts ranked the item as indicating low severity whereas the model placed the item in the mid-range of problem gambling severity.

Discussion

CFA confirmed a one-factor structure is the best fit for the PGSI. The internal consistency was 0.86, comparable to the value reported in the development sample (0.84) (Ferris and Wynne, 2001) and the factor analytic studies by Holtgraves (2009) and Orford *et al.* (2010). The lowest loading overall and across the various subgroups (gender, age, income, type of game) was “borrowing money”, which likely reflects its relatively low endorsement in a general population sample. Importantly, the factor structure was invariant across subgroups defined by gender, age, income level, and game type. Furthermore, the Rasch model applied to the PGSI revealed a well-fitting, unidimensional model overall with no miss-fitting items. The estimated internal consistency of the PGSI was very high (1.00) and standard errors were uniformly low with a small range (0.03–0.05).

Although this type of analysis has not been conducted on the PGSI before, Rasch models of other problem gambling measures indicate some similarities. For instance, three other Rasch studies also found “chasing losses” (as assessed by the SOGS and NODS) to being indicative of low severity levels of problem gambling (Molde *et al.*, 2010; Strong *et al.*, 2004a; Strong and Kahler, 2007). Additionally, we replicated the finding of “financial problems” as a high severity item (Strong and Kahler, 2007). Encouragingly, our results were similar to previous analyses of the PGSI (Holtgraves, 2009; Orford *et al.*, 2010), in that “chasing losses”, “feeling guilty”, and “betting more than one could afford to lose” were low severity items, whereas “borrowing money” and “financial problems” were high severity items. Furthermore, the fact that these low severity items have replicated across three different item analyses of the PGSI supports the efficacy of this instrument as a screening tool for problem gambling risk.

“Borrowing money” was the only PGSI item to display both a statistically and clinically meaningful DIF between game type subgroups, suggesting a bias exists in the item that may result in the severity of problem gambling being under-estimated in EGM and casino gamblers. Specifically, endorsement of the need to borrow money occurs at higher levels of severity for EGM or casino gamblers compared to gamblers who play other games. Borrowing money also had the lowest factor loading. The weaker psychometric properties for this item may be the result of the generally low rate of endorsement in the general population, the ambiguous wording of the item, or both. The present item wording, “Have you borrowed money or sold anything to get money to gamble?” could be subject to different interpretations. The item could be improved by removing the dual question (“borrowed money or sold anything”) and providing examples of what constitutes borrowing.

Int. J. Methods Psychiatr. Res. 22(3): 245–255 (2013). DOI: 10.1002/mpr

There may be phenomenological differences between EMG/casino and other types of gamblers that may also account for the DIF finding. For example, EGM/casino gamblers may engage in more subtle forms of borrowing activities (e.g. using credit card advances) that are under-reported on the item as presently worded. Providing examples of all forms of borrowing may diminish the difference in item endorsement between gambler types. However, before modifying the item it would be important to understand the patterns of borrowing across habitual players of different game types. Relating this data to problem gambling severity would help to inform how the PGSI item on borrowing money could be improved. Qualitative analysis such as cognitive interviewing (a procedure in which subjects are invited to "think aloud" how they respond to items) could also illuminate how different gamblers interpret this item.

Similar to other problem gambling measures, the PGSI items did not provide consistent coverage across the entire problem gambling continuum. Specifically, several items clustered around the low to moderate severity range with a gap between these two points. In a recent study we examined the discriminant validity of all four PGSI interpretative categories: non-problem, low-risk, moderate-risk, and problem gambler (Currie *et al.*, 2013). In terms of discriminant validity, the low-risk and moderate-risk categories were the most problematic. A contributing problem is the limited range of scores within these categories. The validity of these categories was improved with a simple change to the scoring rules (defining low-risk scoring 1–4 and moderate-risk as scoring 5–7). However, the addition of other mid-range severity symptoms would undoubtedly improve the functioning of the PGSI as a continuous measure of severity particularly for gamblers who are below the threshold for problem gambling. New items could increase the range of scores for gamblers classified as low-risk or moderate-risk (McCready and Adlaf, 2006). In fact, several non-scored items have been included in versions of the CPGI administered in population surveys. For example, the CPGI used in the CCHS-1.2 included questions on conflict with family or friends, hiding gambling activities from others, and gambling to forget problems (Statistics Canada, 2002).

Items that may better identify at-risk gamblers are not limited to self-reported symptoms. Quantitative indicators of gambling intensity (e.g. exceeding low-risk gambling thresholds for frequency and amount of money spent on gambling; Currie *et al.*, 2008b) could also be used as signs of at-risk gambling. Recent longitudinal data revealed that playing EGM or casino games at least monthly predicts the progression from low-risk to high-risk gambling (Currie *et al.*, 2011). Given that other prominent measures within this field face similar difficulties, it is encumbant for future research to investigate potential symptoms that may emerge as low-risk problem gambling escalates.

It is noteworthy that the Rasch generated severity estimates closely corresponded with expert rankings of symptom severity, supporting the ecological validity of this instrument. In particular, experts rated "health problems", "financial problems", and "borrowing money" as reflective of high problem gambling severity; a result parallel to Rasch modelling results. Similarly, "recognizing problem with gambling" and "tolerance" were indicative of moderately severe problem gambling, and "feelings of guilt" represented low levels of problem gambling according to expert ratings and Rasch modelling. For three items, there were discrepancies between expert ratings and Rasch modelling, however the correlation between expert ratings and the Rasch model estimates remained quite high ($r > 0.80$). We acknowledge several limitations with this data. Respondents to our survey were provided no a priori information on the expected ranking of severity and hence used their own experience and intuitive judgments to rate individual items. We fully expect that experts aligned their ratings with personally held theories of problem gambling. This may explain the high severity rating given chasing losses, a symptom that is synonymous with pathological gambling in most theoretical frameworks. We feel our approach was justified for several reasons. Foremost, we did not want to influence experts' ratings. Similarly, the Rasch model is not a test of specific severity ratings rather the estimates are allowed to emerge from the patterns in the data. We also surveyed experts with a variety of professional backgrounds and roles in the field of gambling (researchers, clinicians and regulators) who undoubtedly represent a variety of theoretical perspectives on the phenomenology of gambling. The ratings therefore represent the average of these theoretical frameworks.

This study was limited because the PGSI was the only measure of gambling problems, therefore we could not assess DIF for non-problem gamblers versus problem gamblers. Nonetheless, type of game may be a gross indicator of problem gambling (i.e., VLT/casino gamblers tend to have more severe gambling problems compared to other types of gamblers. However, future studies should examine the DIF of the PGSI using both community samples of gamblers and clinical samples of pathological gamblers). Furthermore, future studies should address the PGSI's gaps in coverage across the problem gambling continuum.

Acknowledgements

This research was funded by research grants from the Alberta Gaming Research Institute and the Canadian Consortium for Gambling Research (consisting of Alberta Gaming Research Institute, Canadian Centre on Substance Abuse, Canadian Partnership for Responsible Gambling, Manitoba Gaming Control Commission, Nova Scotia Gaming Foundation, Ontario Problem Gambling Research Centre, Saskatchewan Health). The authors also thank Lynda Edwards and Lynda Barrons who helped in manuscript preparation.

References

Abbott M.W., Volberg R. (2006) The measurement of adult problem and pathological gambling. *International Gambling Studies*, **6**(2), 175–200, DOI: 10.1080/14459790600928678

Bond T.G., Fox C.M. (2007) *Applying the Rasch Model. Fundamental Measurement in the Human Sciences*, 2nd edition, Mahwah, NJ, Lawrence Erlbaum Associates Publishers.

Breen R., Zimmerman M. (2002) Rapid onset of pathological gambling in machine gamblers. *Journal of Gambling Studies*, **18**(1), 31–43, DOI: 10.1023/A:1014580112648

Cox B.J., Yu N., Afifi T.O., Ladouceur R. (2005) A national survey of gambling problems in Canada. *Canadian Journal of Psychiatry – Revue Canadienne de Psychiatrie*, **50**(4), 213–217.

Currie S.R., Hodgins D.C., Casey D. (2013) Validity of the problem gambling severity index interpretive categories. *Journal of Gambling Studies*, **29**(2), 311–327, DOI: 10.1007/S10899-012-9300-6

Currie S.R., Hodgins D.C., Casey D., el-Guebaly N., Smith G., Williams R.J., Schopflocher D., Wood R. (2011) Examining the predictive validity of low risk gambling limits with longitudinal data. *Addiction*, **107**(2), 400–406, DOI: 10.1111/j.1360-0443.2011.03622.x

Currie S.R., Hodgins D.C., el-Guebaly N., Manley C., Coulson R. (2004) Factor validation of the Addiction Severity Index Scale structure in persons with concurrent disorders. *Psychological Assessment*, **16**(3), 326–329, DOI: 10.1037/1040-3590.16.3.326

Currie S.R., Hodgins D.C., Wang J., el-Guebaly N., Wynne H. (2008a) In pursuit of empirically based responsible gambling limits. *International Gambling Studies*, **8**(2), 209–229, DOI: 10.1080/14459790802172265

Currie S.R., Hodgins D.C., Wang J., el-Guebaly N., Wynne H., Chen S. (2006) Risk of harm from gambling in the general population as a function of level of participation in gambling activities. *Addiction*, **101**(4), 570–580, DOI: 10.1111/j.1360-0443.2006.01392.x

Currie S.R., Hodgins D.C., Wang J., el-Guebaly N., Wynne H., Miller N.V. (2008b) Replication of low-risk gambling limits using Canadian provincial gambling prevalence data. *Journal of Gambling Studies*, **24**(3), 321–335, DOI: 10.1007/s10899-008-9091-y

Currie S.R., Miller N.V., Hodgins D.C., Wang J. (2008c) Defining a threshold of harm from gambling for population health surveillance research. *International Gambling Studies*, **9**(1), 19–38, DOI: 10.1080/14459790802652209

Department of Social Development (Northern Ireland). (2010) *Northern Ireland Gambling Prevalence Survey 2010 – Statistical First Release*, Belfast: Department of Social Development.

Embretson S.E., Reise S.P. (2000) *Item Response Theory for Psychologists*, Mahwah, NJ, Erlbaum.

Ferris J., Wynne H.J. (2001) *The Canadian Problem Gambling Index Final Report*. Ottawa, ON: Canadian Centre on Substance Abuse.

Garrett A.M. (2003) Rasch analysis of the Roland Disability Questionnaire. *Spine*, **28**(1), 79–84.

Gerstein D.R., Volberg R.A., Toce M.T., Harwood H., Johnson R.A., Buie T., Christiansen E. (1999) *Gambling Impact and Behavior Study: Report to the National Gambling Impact Study Commission*, Chicago, IL: NORC, at the University of Chicago.

Gravel R., Beland Y. (2005) The Canadian Community Health Survey: Mental Health and Well-Being. *Canadian Journal of Psychiatry – Revue Canadienne de Psychiatrie*, **50**(10), 573–579.

Holtgraves T. (2009) Evaluating the Problem Gambling Severity Index. *Journal of Gambling Studies*, **25**(1), 105–120, DOI: 10.1007/s10899-008-9107-7

Hoyle R.H., Smith G. (1994) Formulating clinical research hypotheses as structural equation models: a conceptual overview. *Journal of Consulting & Clinical Psychology*, **62**(3), 429–440.

Hu L.T., Bentler P.M. (1999) Cutoff criteria for fit indices in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling*, **6**(1), 1–55, DOI: 10.1080/10705519909540118

Kavli H., Berntsen W. (2005) *Gambling Habits and Gambling Problems in the Population. Prepared for Norsk Tipping, the Government Gambling Operator*, Olso: MMI Research.

Kristiansen S., Jensen S. (2011) Prevalence of gambling problems among adolescents in the Nordic countries: an overview of national gambling surveys 1997–2009. *International Journal of Social Welfare*, **20**(1), 75–86, DOI: 10.1111/j.1468-2397.2009.00701.x

Langenbucher J.W., Labouvie E., Martin C.S., Sanjuan P.M., Bavly L., Kirisci L., Chung T. (2004) An application of item response theory analysis to alcohol, cannabis, and cocaine criteria in DSM-IV. *Journal of Abnormal Psychology*, **113**(1), 72–80, DOI: 10.1037/0021-843X.113.1.72

Lesieur H., Blume S. (1987) The South Oaks Gambling Screen (SOGS): a new instrument for the identification of pathological gamblers. *American Journal of Psychiatry*, **144**(9), 1184–1188.

Linacre J.M. (1999) Category disordering vs. step (threshold) disordering. *Rasch Measurement Transactions*, **13**, 675.

Linacre J.M. (2003) Rasch power analysis: size vs. significance: standardized chi-square fit statistics. *Rasch Measurement Transactions*, **17**(1), 918.

McCready J., Adlaf E. (2006) *Performance and Enhancement of the Canadian Problem Gambling Index: Report and Recommendations*, Guelph: ON, Ontario Problem Gambling Research Centre.

Mcmillen J., Wenzel M. (2006) Measuring problem gambling: assessment of three prevalence screens. *International Gambling Studies*, **6**(2), 147–174, DOI: 10.1080/14459790600927845

Miller N.V., Currie S.R. (2008) A Canadian population level analysis of the roles of irrational gambling cognitions and risky gambling practices as correlates of gambling intensity and pathological gambling. *Journal of Gambling Studies*, **24**(3), 257–274, DOI: 10.1007/s10899-008-9089-5

Ministry of Health. (2009) *A Focus on Problem Gambling: Results of the 2006/07 New Zealand Health Survey*, Wellington: Ministry of Health.

Molde H., Hystad S.W., Pallesen S., Myrseth H., Lund L. (2010) Evaluating lifetime NODS using Rasch modelling. *International Gambling*

Studies, **10**(2), 189–202, DOI: 10.1080/14459795.2010.502182

Neal P., Delfabbro P., O'Neil M. (2005) *Problem Gambling and Harm: Towards a National Definition*, Melbourne, Gambling Research Australia.

Olason D., Gretarsson S. (2009) Iceland. In Meyer G., Hayer T., Griffiths M. (eds) *Problem Gambling in Europe: Challenges, Prevention, and Interventions*, pp. 137–152, New York, Springer.

Orford J., Wardle H., Griffith 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, DOI: 10.1080/14459790903567132

Smith E.V.J., Wakely M.B., De Kruif R.E.L., Swartz C.W. (2003) Optimizing rating scales for self-efficacy (and other) research. *Educational and Psychological Measurement*, **63**(3), 369–391, DOI: 10.1177/0013164403063003002

Statistics Canada. (2002) *Canadian Community Health Survey (Cycle 1.2) – Mental Health and Well-being User's Guide*, Ottawa, ON: Statistics Canada.

Steinberg L., Thissen D.E.-M.A. (2006) Using effect sizes for research reporting: examples using item response theory to analyze differential item functioning. *Psychological Methods*, **11**(4), 402–415, DOI: 10.1037/1082-989X.11.4.402

Stinchfield R. (2002) Reliability, validity, and classification accuracy of the South Oaks Gambling Screen (SOGS). *Addictive Behaviors*, **27**(1), 1–19, DOI: 10.1016/S0306-4603(00)00158-1

Stinchfield R., Govoni R., Frisch G.R. (2007) A review of screening and assessment instruments for problem and pathological gambling. In Smith G., Hodgins D.C., Williams R.J. (eds) *Research and Measurement Issues in Gambling Studies*, pp. 179–213, Burlington, MA, Academic Press.

Strong D.R., Breen R.B., Lesieur H.R., Lejuez C.W. (2003) Using the Rasch model to evaluate the South Oaks Gambling Screen for use with nonpathological gamblers. *Addictive Behaviors*, **28**(8), 1465–1472, DOI: 10.1016/S0306-4603(02)00262-9

Strong D.R., Breen R.B., Lejuez C.W., Strong D.R. (2004a) Using item response theory to examine gambling attitudes and beliefs. *Personality & Individual Differences*, **36**(7), 1515–1529, DOI: 10.1016/j.paid.2003.06.001

Strong D.R., Kahler C.W. (2007) Evaluation of the continuum of gambling problems using the DSM-IV. *Addiction*, **102**(5), 713–721, DOI: 10.1111/j.1360-0443.2007.01789.x

Strong D.R., Lesieur H.R., Breen R.B., Stinchfield R., Lejuez C.W. (2004b) Using a Rasch model to examine the utility of the South Oaks Gambling Screen across clinical and community samples. *Addictive Behaviors*, **29**(3), 465–481, DOI: 10.1081/JA-120030897

Volberg R., Bernhard B. (2006) *The 2006 Study of Gambling and Problem Gambling in New Mexico*. Northampton: MA, Gemini Research.

Wardle H., Sproston K., Orford J., Erens B., Griffiths M., Constantine R. (2007) *British Gambling Prevalence Survey 2007*, London: National Centre for Social Research.

Welte J., Barnes G., Wieczorek W., Tidwell M.-C., Parker J. (2004) Risk factors for pathological gambling. *Addictive Behaviors*, **29**(2), 323–335, DOI: 10.1016/j.addbeh.2003.08.007

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

Wright B.D., Stone M.H. (1979) *Best Test Design: Rasch Measurement*, Chicago, IL, MESA Press.

Young M., Stevens M. (2008) SOGS and CPGI: parallel comparison on a diverse population. *Journal of Gambling Studies*, **24**(3), 337–356, DOI: 10.1007/s10899-007-9087-z