

ORIGINAL ARTICLE

Development and validation of the Gambling Follow-up Scale, Self-Report version: an outcome measure in the treatment of pathological gambling

Ana M. Galetti, Hermano Tavares

Programa Ambulatorial do Jogo, Instituto e Departamento de Psiquiatria, Faculdade de Medicina, Universidade de São Paulo (USP), São Paulo, SP, Brazil.

Objective: To validate the Gambling Follow-up Scale, Self-Report version (GFS-SR), a 10-item scale designed to assess gambling frequency, time and money spent on gambling, gambling craving, debts, emotional distress, family relationships, autonomy, and frequency of and satisfaction with leisure activities in individuals diagnosed with gambling disorder according to the DSM-5 criteria.

Methods: One hundred and twenty treatment-seeking gamblers were evaluated, 84 of whom proceeded to treatment. Fifty-two relatives provided collateral informant reports at baseline. Six months later, the 50 patients who completed the program were reassessed.

Results: The GFS-SR showed good inter-rater agreement and internal consistency. Factor analysis presented a three-factor solution: gambling behavior (factor 1); social life (factor 2); and personal hardship (factor 3). There was a high degree of convergence between GFS-SR scores and those of reference scales. The GFS-SR scores showed excellent sensitivity to change (factor 1), predictive validity for treatment response (factor 2), and ability to distinguish recovered from unrecovered patients after treatment (factor 3). A cutoff score of 33 was found to have 87% sensitivity and 80% specificity for gambling recovery.

Conclusion: The GFS-SR is well suited to providing reliable follow-up of gamblers under treatment and assessing the efficacy of their treatment.

Keywords: Gambling; treatment outcome; questionnaires; psychometrics; self-report

Introduction

Gambling has become an international concern in recent years. The prevalence of gambling disorder (GD) has risen, as the number of individuals exposed to a variety of games of chance has increased.¹ However, less than 10% of gamblers seek help. Furthermore, the majority of gamblers recover without assistance.² Evidently, the field of gambling research is in need of a flexible method for reliable assessment of recovery in different contexts, such as formal treatment, lay intervention, self-help, and natural recovery.

Recent efforts have been made to establish universal goals for gambling treatment and remission. In the Banff Consensus,³ a panel of experts proposed a minimum set of indicators for reporting treatment outcomes, divided into three domains: measures of gambling behavior (monthly expenditure on gambling, number of days gambling per month, and time spent on gambling-related pursuits); measures of the problems caused by gambling (personal health, relationship, financial, and legal problems), which can be supplemented by quality of life measures; and

measures of the process of change specifically related to the treatment modality applied (e.g., measures of cognitive distortions about random events in the case of cognitive restructuring therapy for GD). However, despite efforts to achieve a consensus for reporting treatment outcomes, there is as yet no standard indicator of the efficacy of gambling treatment. Pharmacological trials have placed emphasis on craving and related phenomena in gambling.⁴ Reports on controlled psychosocial interventions focus mostly on indicators of gambling behavior such as gambling frequency, amount of money spent on gambling, perception of control, and perceived self-efficacy.⁵⁻⁹ This heterogeneity of outcome measures hampers comparisons between and among treatment modalities.

Various scales have been developed with the aim of diagnosing GD and assessing gambling severity. The South Oaks Gambling Screen (SOGS) is probably the most traditional and widely used instrument in the field.¹⁰ It can be used in an interview or self-report format. As a diagnostic instrument, the SOGS offers good to excellent specificity and score reliability. However, it is based upon criteria from the third edition of the DSM, which, over the last three decades, has undergone four reviews, until the publication of the current version of the manual, the DSM-5. The DSM-III criteria were strongly criticized for their heavy emphasis on money problems caused by gambling, and the same criticism applies to the SOGS.¹¹

Correspondence: Hermano Tavares, Instituto de Psiquiatria, Faculdade de Medicina, Rua Dr. Ovídio Pires de Campos, 785, 1º andar, sala 4, CEP 05403-010, São Paulo, SP, Brazil.
E-mail: hermanot@uol.com.br

Submitted Jan 07 2016, accepted May 11 2016.

In addition, the SOGS was originally designed as a screening tool; its first version was meant to assess lifetime problems with gambling. A later version of the SOGS was adapted for a 12-month reference period, which preserved its psychometric strength, but failed to accurately distinguish between the different degrees of gambling severity. This adds to the perception that screening and diagnostic instruments may not automatically translate into good severity assessment tools for the follow-up of gamblers under treatment, who may require periodic evaluation over shorter time intervals.¹² The Gambling Symptom Assessment Scale (G-SAS), created by Kim et al., is a 12-item self-report instrument developed for the follow-up of patients undergoing trials of pharmacological therapy for GD. It taps both into subjective gambling experiences (urges, thoughts, and anticipation) and the objective aspects of gambling behavior. The G-SAS presented good to excellent psychometric properties, as well as good cross-validity with the Clinical Global Impression.¹³ Nonetheless, the focus on gambling consequences is too narrow, with one item for the assessment of emotional distress and another clustering relationship, financial, legal, job, and health problems. Therefore, it fails to provide a broader assessment of gambling-related distress and may not be entirely suitable for patients undergoing nonpharmacological interventions, or for individuals undergoing a natural recovery process.

In the DSM-IV-TR, GD was designated pathological gambling and listed among the impulse control disorders.¹⁴ In the DSM-5, however, GD was reclassified, and is now listed in a new section titled addictive disorders.^{15,16} In fact, given the similarities between GD and substance use disorders, one natural strategy for development of instruments to assess gambling has always been to adapt tools from the addiction field. The Addiction Severity Index (ASI), which is widely used in this area,¹⁷ is a semi-structured interview that assesses seven variables: medical condition; employment or support; alcohol use; drug use; illegal activity; family and social relationships; and psychiatric condition. Petry^{18,19} validated a five-question gambling section for the ASI. Altogether, the ASI sections provide a broad assessment of gamblers under treatment and, unlike the scales reported above, are not limited to assessing gambling behavior. However, the ASI involves an extensive interview and must be administered by an interviewer who has received specific training, which can be an obstacle when swift and repeated measures are needed. Finally, the timeline follow-back (TLFB) interview is a method initially developed to reliably retrieve and assess information on alcohol consumption over a specific period of time.²⁰ Hodgins & Makarchuk²¹ used TLFB interviews to pinpoint gambling behavior, a strategy further validated by Weinstock et al.²² The TLFB interview adapted for GD (TLFB-GD) provides a retrospective diary of gambling activity with excellent test-retest reliability and strong correlation with collateral informant reports.¹⁸

Clearly, the instruments and methods explained above have complementary natures, covering most of the topics proposed in the Banff Consensus.² However, applying all

of them in a single investigation would consume more time than is feasible. In addition, to our knowledge, studies employing such scales have not produced criteria that enable clinicians to determine GD remission. Therefore, integrating these tools into a single, reliable, rapidly scored instrument would be helpful. Establishing a reliable and simple assessment of gambling remission could benefit not only researchers but also clinicians and other non-professional personnel involved in other types of care, who are frequently overburdened by the need to control several comorbid conditions and a multitude of related challenges in addition to GD. With this goal in mind, our group developed and tested an initial version of the Gambling Follow-up Scale (GFS), with the objective of providing measures that would be widely accepted as standards for the assessment of gambling treatment.²³ The GFS has five domains: frequency of and time spent gambling; work; family relationships; leisure; and Gamblers Anonymous (GA) attendance. The last domain was added to determine whether membership in GA would provide additional therapeutic support. The purpose of our previous validation study was to ensure the suitability of the GFS for assessing as many different gambling treatments as possible, easily combining with measures of the change process related to the treatment modality. As a semi-structured interview, the GFS can be administered in approximately 6 min. In addition, the GFS displayed excellent reliability, with inter-rater agreement ranging from 82 to 95%, and intraclass coefficients ranging from 0.85 to 0.99 (all $p < 0.001$).

With the general aim of further developing the GFS into a broader, yet still rapidly applied, self-report version, the GFS-SR (Appendix 1, online-only supplementary material), our primary goals in the present study were to investigate the psychometric properties of the GFS-SR (scoring convergence with reference scales, inter-rater agreement, internal consistency, factorial structure, and sensitivity to change) and to establish a cutoff GFS-SR score that would reliably indicate GD remission (i.e., the patient no longer meets the DSM-5 criteria for pathological gambling). A secondary goal was to explore predictors of gambling remission among the measures obtained from the GFS-SR and from the other gambling scales used in its cross-validation.

Methods

Participants

We evaluated 120 consecutive patients who sought treatment at the Gambling Outpatient Unit of Instituto e Departamento de Psiquiatria, Faculdade de Medicina, Universidade de São Paulo, in the city of São Paulo, Brazil, between November 2002 and September 2004. All 120 patients met the DSM-IV-TR criteria for a diagnosis of pathological gambling, and none refused to participate in the study. Of the 120 subjects initially assessed, 84 (70%) voluntarily proceeded to treatment. Patients were invited to bring one relative to provide collateral information; 52 such relatives completed a version of the GFS-SR adapted for collateral informants (Appendix 2, online-only

supplementary material). Of the 84 patients who entered treatment, 50 (60%) completed the program. It is important to note that data collection occurred prior to the publication of the DSM-5. The criteria for pathological gambling changed in the DSM-5, but, fortunately, the DSM-IV-TR criteria (10 criteria) can be adapted to the DSM-5 criteria (9 criteria) simply by excluding the criterion referring to “illegal acts” to sustain gambling, thus reducing the cutoff from five to four criteria.¹⁵

All work with human subjects reported in this study complied with the guidelines and principles for experimental procedures of the Helsinki Declaration.²⁴ The study was approved by the Research Ethics Committee of Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, and all participants gave written informed consent.

Scoring

Trained psychologists and psychiatrists specializing in pathological gambling evaluated all of the patients. The following instruments were applied: GFS-SR; gambling subscale of the ASI (ASI-G); TLFB-GD; G-SAS; and Social Adjustment Scale.

The GFS-SR total score is the sum of all items. All items are scored from 1 to 5, with two exceptions: item 5, a multiple-choice question with only four choices and therefore scored from 1 to 4; and item 9, a multi-part question related to the frequency of leisure activities, with each part being scored from 0 to 4. We tested two ways of scoring item 5: adding its raw score to the total GFS-SR score; and adjusting its values to match a five-point scale (i.e., 1 point on the four-point scale would correspond to 1.25 points in the total score). The statistical analyses conducted for both scoring schemes yielded quite similar outcomes. Therefore, we adopted the first scoring scheme. The structure of item nine required standardization to adjust its original range of scores (0-48) to that of the other items in the GFS-SR (1-5). The final score of item 9 is thus calculated as the sum of all frequencies of leisure activities (activities not selected, including the other option, are scored zero), divided by 12, plus one, as shown in the formula below:

$$\text{item 9 score} = \left(\sum \text{sub-item scores} / 12 \right) + 1$$

Score reliability, internal consistency, and factor analysis

We tested the consistency of patient reports at baseline by measuring the agreement between self-reports and collateral informant reports (n=52 pairs), using Cohen's kappa. We interpreted Cohen's kappa values as recommended by Landis & Koch, considering values < 0.2 as indicative of poor agreement, 0.2-0.4 as indicative of fair agreement, and > 0.4 as indicative of better (moderate to very good) agreement.²⁵ The GFS-SR adapted for the collateral informant provides information about gambling frequency, time spent gambling, money spent on gambling, family relationships, autonomy, and frequency of leisure activities (corresponding to GFS-SR items 1, 2, 3,

7, 8, and 9, respectively). The other items were considered subjective in nature and difficult for an external observer to assess. The reliability of the GFS-SR scoring was assessed by Cronbach's alpha, and the data related to the 120 treatment-seeking gamblers initially evaluated were processed via exploratory factor analysis to investigate the structure of the GFS-SR.

Convergent validity

The ASI-G questions used in this study were the number of days on which gambling occurred, money spent on gambling, and the number of days of worry due to gambling, all relating to the last 4 weeks. The TLFB-GD domains evaluated were the number of days on which gambling occurred, total hours spent gambling, and money spent on gambling, all in the last 4 weeks. The Social Adjustment Scale is a 54-item, self-report scale that assesses seven specific areas: work (external, domestic, or academic); social life and leisure; family relationships (with parents, siblings, and other relatives); marital relationship; relationship with children; domestic life; and financial situation. It has shown sensitivity to distinguish individuals with depression, alcohol use disorder, or schizophrenia from healthy controls, as well as displaying sensitivity to change in psychological and pharmacological trials.²⁶ The Portuguese-language version of the Social Adjustment Scale was validated by Gorenstein et al.²⁷ by comparing healthy individuals to depression, panic, cocaine abuse, and bulimic patients (n=174). The Social Adjustment Scale scores for healthy individuals and depressed patients were similar to those reported in previous studies using its original version. The Cronbach's alpha coefficient for the subscales was 0.85; healthy individuals scored lower than psychiatric patients (all p < 0.001). The Social Adjustment Scale also displayed ability to discriminate between patients in remission and those in acute panic or depressed patients (p-values ≤ 0.006), and correlated significantly with the Beck Depression Inventory and the Clinical Global Impression scale.²⁷

The G-SAS was translated into Portuguese by the authors, with specific attention to the cultural equivalence of terms related to gambling.²⁸ While conducting the GFS-SR validation study, we availed ourselves of this opportunity to check the psychometric properties of the Portuguese version of the G-SAS. Cronbach's alpha coefficient for the Portuguese G-SAS was 0.923. A factorial analysis found a three-factor solution, with one factor grouping items from 1 to 8 (gambling urges, thoughts, and behavior), explaining 45% of the variance; a second factor grouping items 9 and 10 (emotional anticipation); and a third factor grouping items 11 and 12 (gambling-related harm), accounting for 16% and 15% of the variance, respectively. The Portuguese G-SAS score correlated significantly with the ASI-G and TLF-GD measures of gambling frequency and money spent on gambling, respectively, as well as with the Portuguese Social Adjustment Scale (Spearman's $\rho = -0.299$ to -0.824 , n=120, all significant at p < 0.001). Further detail on the psychometric performance of the Portuguese G-SAS

would exceed the scope of the present paper, and will be the subject of a future publication.

We also explored associations between the DSM-5 criteria and the GFS-SR. The DSM-5 establishes three severity levels for GD (four or five criteria, mild; six or seven criteria, moderate; eight or nine criteria, severe).¹⁵ We compared these severity levels to the GFS-SR item partial and total scores using the Kruskal-Wallis test and the Mann-Whitney *U* test for post-hoc comparisons between categories. Finally, we ran Spearman's correlation analyses between the total number of DSM-5 criteria met and the GFS-SR scores at the beginning and end of treatment.

Sensitivity to change over time

The patients entered a 6-month program (intervention) that comprised a medical assessment, aimed at diagnosis and treatment of psychiatric comorbidities,²⁹ and a brief cognitive intervention targeting cognitive distortions regarding randomness and games of chance.³⁰ Although GA attendance was not mandatory, it was encouraged. Fifty patients were reassessed at treatment completion with the same scales used at baseline. We compared pre- and post-treatment scores using the Wilcoxon test. Item-by-item analyses were conducted for the GFS-SR, ASI-G, and TLFB-GD, because the latter two did not yield total scores.

Discriminant validity, outcome prediction, and cutoff score

To analyze the ability of the GFS-SR to discriminate between responders and nonresponders (i.e., to identify gambling remission), we divided the sample according to the number of DSM-5 GD criteria met at the end of treatment. Patients meeting fewer than four of the criteria were classified as recovered (responders), whereas those meeting four or more were classified as unrecovered (nonresponders).

To adjust for differences in baseline status, we subtracted the initial score of each scale from the corresponding final score and used the absolute positive value as the variation index of each scale. We then used the Mann-Whitney *U* test to compare the variation scores between recovered and unrecovered patients. Finally, the variation indices that reached significance ($p < 0.050$) in this preliminary univariate analysis were entered into a logistic regression model in which recovery from GD was the dependent variable.

Additional analyses were conducted to account for the effects of parallel GA attendance. The potential predictive value of the pre-treatment assessment measures was tested by comparing recovered and unrecovered patients in terms of demographic profile and gambling variables at baseline. Then, we applied a receiver operating characteristic (ROC) curve analysis, using the above-mentioned DSM-5 criteria for gambling recovery and the GFS-SR score at the end of treatment as the test variables. Finally, we ran an alternative data analysis using the DSM-IV-TR criteria for pathological gambling, but since the patients classified as responders and nonresponders remained the same, the outcomes obtained were quite similar to those presented in the Results section. Data analysis was performed in SPSS version 8.0.³¹

Results

Sample

Among the 120 subjects initially assessed, the mean age was 44.1 (standard deviation [SD] = 10.3) years; the mean number of years of education was 12.2 (SD = 3.9); 76% were white; 60% were married; and 51% were female. In 82%, electronic slot machines were the preferred form of gambling. As mentioned previously, 84 (70%) of the eligible subjects opted for treatment. Of those 84 subjects, 50 (60%) completed the treatment program. We found no significant differences in demographic profile on comparison of subjects who opted out of treatment ($n=36$), those who dropped out of treatment ($n=34$), and those who completed treatment ($n=50$). In addition, we found no statistical differences between the patients who completed the program ($n=50$) and those who did not ($n=70$), in terms of pretreatment G-SAS, Social Adjustment Scale, and GFS-SR scores (Mann-Whitney *U* ranging from -0.194 to -1.087, p ranging from = 0.227 to 0.846).

Collateral agreement, score reliability, and factor analysis

The agreement between self-reports and collateral informant reports ranged from fair to moderate, with kappa coefficients ranging from 0.241 to 0.486 (p ranging from ranging from 0.001 to < 0.001 , $n=52$ pairs), except for the question concerning family relationships, which presented poor agreement (kappa = 0.170, $p = 0.039$, $n=52$ pairs) (Table 1). Cronbach's alpha coefficient for the GFS-SR as a whole was 0.855 ($n=120$), and all items correlated well with the scale as a whole (Table 2).

Factor analysis based on the initial sample ($n=120$) presented a three-factor solution (Table 3), in which 68.9% of the variance in GFS-SR scores was explained. Factor 1 (designated gambling behavior) comprised the items gambling frequency, time spent gambling, money spent on gambling, and gambling craving; factor 2 (designated social life) comprised the items family relationships, frequency of leisure activities, and satisfaction with leisure activities; and factor 3 (designated personal hardship) comprised the items debts, emotional distress, and autonomy.

Table 1 Agreement between self-reports and collateral informant reports ($n=52$ pairs)

GFS-SR item vs. GFS-CI item	Patients vs. collateral informants	
	kappa	p-value
Gambling frequency	0.409*	< 0.001
Time spent gambling	0.486*	< 0.001
Money spent on gambling	0.356 [†]	< 0.001
Family relationships	0.170 [‡]	0.039
Autonomy	0.297 [†]	< 0.001
Frequency of leisure activities	0.214 [†]	0.001

GFS-CI = Gambling Follow-up Scale, Collateral Informant version; GFS-SR = Gambling Follow-up Scale, Self-Report version.

* Moderate, [†] fair, and [‡] poor agreement.

Convergent validity

The average time to completion of the GFS-SR was 6 min. In general, GFS-SR items had moderate to excellent cross-validity with their corresponding scoring on

Table 2 Internal consistency of the Gambling Follow-up Scale, Self-Report version (n=120)

GFS-SR item (description)	Item-total correlation	Alpha if item deleted*
1 (gambling frequency)	0.696	0.829
2 (time spent gambling)	0.691	0.830
3 (money spent on gambling)	0.705	0.828
4 (gambling craving)	0.679	0.831
5 (debts)	0.595	0.841
6 (emotional distress)	0.565	0.842
7 (family relationships)	0.418	0.853
8 (autonomy)	0.462	0.853
9 (frequency of leisure activities)	0.486	0.853
10 (satisfaction with leisure activities)	0.422	0.854

GFS-SR = Gambling Follow-up Scale, Self-Report version.

*Cronbach's alpha for the GFS-SR as a whole (all items included) = 0.855.

the reference scales (Spearman's $\rho = -0.299$ to -0.824 , $n=120$, all significant) (Table 4). In addition, the GFS-SR total score correlated significantly with the G-SAS total score ($\rho = -0.796$, $p < 0.001$), the Social Adjustment Scale total score ($\rho = -0.645$, $p < 0.001$), TLFB-GD hours spent on gambling ($\rho = -0.751$, $p < 0.001$), and TLFB-GD money spent on gambling ($\rho = -0.645$, $p < 0.001$).

At the start of treatment ($n=120$), six patients (5%) were classified as having mild GD, and 39 (32.5%) and 75 (62.5%) as having moderate and severe GD, respectively. At this point, only the GFS-SR factor 2 (social life) was found to be associated with the DSM-5 severity categories ($\chi^2_{(2)} = 10.18$, $p = 0.006$), with moderate gamblers scoring higher (mean [M] = 8.1, SD = 3.0) than did severe gamblers (M = 6.5, SD = 2.8, $U = 987.0$, $p = 0.004$). Individually, only items 5 (debts, $p = 0.039$) and 8 (autonomy, $p = 0.002$) significantly differentiated moderate from severe gamblers. The total number of DSM-5 GD criteria met did not correlate with the G-SAS or TLFB measures, correlated only marginally with the GFS-SR total score ($\rho = -0.163$, $p = 0.080$), and correlated significantly with factor 2 (social life) ($\rho = -0.306$, $p = 0.001$).

Table 3 Factor analysis of the Gambling Follow-up Scale, Self-Report version (n=120)

GFS-SR item (description)	Factor loadings*†‡§		
	Factor 1 (VE = 29.3%)	Factor 2 (VE = 19.9%)	Factor 3 (VE = 19.7%)
1 (gambling frequency)	0.875	0.185	0.130
2 (time spent gambling)	<u>0.729</u>	0.285	0.232
3 (money spent on gambling)	<u>0.710</u>	0.128	0.417
4 (gambling craving)	<u>0.804</u>	0.209	0.171
5 (debts)	<u>0.370</u>	0.082	<u>0.738</u>
6 (emotional distress)	0.457	0.042	<u>0.581</u>
7 (family relationships)	0.319	<u>0.635</u>	<u>-0.026</u>
8 (autonomy)	0.051	<u>0.194</u>	<u>0.868</u>
9 (frequency of leisure activities)	0.143	<u>0.811</u>	<u>0.189</u>
10 (satisfaction with leisure activities)	0.107	<u>0.842</u>	0.146

Underlined font indicates the highest factor loading for each item.

GFS-SR = Gambling Follow-up Scale, Self-Report version; VE = variance explained (proportion).

*Varimax rotation with factor extraction at an eigenvalue ≥ 1.00 .

†Kaiser-Meyer-Olkin measure of sampling adequacy = 0.848.

‡Bartlett's sphericity test, approximate $\chi^2_{(45)} = 489.06$, $p < 0.001$.

§Accumulated VE = 68.9%.

Table 4 Correlations between items on the Gambling Follow-up Scale, Self-Report version and questions from reference instruments, in the initial sample (n=120)

GFS-SR item (description)	Reference instrument - correlated question(s)	ρ	p-value
1 (gambling frequency)	ASI-G - days of gambling	-0.824	< 0.001
	TLFB-GD - days of gambling	-0.818	< 0.001
2 (time spent gambling)	TLFB-GD - time spent gambling	-0.816	< 0.001
	ASI-G - money spent	-0.665	< 0.001
3 (money spent on gambling)	TLFB-GD - money spent	-0.668	< 0.001
	SAS - financial situation	-0.414	< 0.001
4 (gambling craving)	G-SAS total score	-0.788	< 0.001
	SAS - financial situation	-0.515	< 0.001
5 (debts)	ASI-G - days of worry	-0.569	< 0.001
6 (emotional distress)	SAS - family relationships	-0.299	0.003
7 (family relationships)	SAS - work	-0.318	< 0.001
8 (autonomy)	SAS - leisure	-0.561	< 0.001
9 (frequency of leisure activities)	SAS - leisure	-0.561	< 0.001
10 (satisfaction with leisure activities)	SAS - leisure	-0.419	< 0.001

ASI-G = gambling subscale of the Addiction Severity Index; GFS-SR = Gambling Follow-up Scale, Self-Report version; G-SAS = Gambling Symptom Assessment Scale; SAS = Social Adjustment Scale; TLFB-GD = timeline follow-back interview adapted for gambling disorder.

Table 5 Sensitivity to change, pre- and post-treatment scores (Wilcoxon test, n=50)

Variable	Pre-treatment	Post-treatment	z	p-value
Number of DSM-5 criteria for GD met	7.6 (1.5)	3.3 (2.8)	-5.54	< 0.001
G-SAS total score	28.5 (10.4)	17.1 (12.1)	-4.74	< 0.001
SAS total score	2.63 (0.66)	2.30 (0.73)	-3.40	0.001
GFS-SR				
Total score	26.9 (7.5)	33.6 (7.7)	-4.44	< 0.001
Factor 1*	12.4 (4.8)	16.2 (4.2)	-4.35	< 0.001
Factor 2 [†]	7.6 (2.4)	8.7 (2.3)	3.04	0.002
Factor 3 [‡]	7.0 (2.2)	8.7 (2.8)	-4.04	< 0.001
Item (description)				
1 (gambling frequency)	3.0 (1.4)	4.3 (1.1)	-4.49	< 0.001
2 (time spent gambling)	3.5 (1.2)	4.3 (1.1)	-4.12	< 0.001
3 (money spent on gambling)	3.1 (1.5)	4.2 (1.4)	-3.52	< 0.001
4 (gambling craving)	2.6 (1.4)	3.5 (1.2)	-3.46	0.001
5 (debts)	2.0 (1.0)	2.5 (0.93)	-3.30	0.001
6 (emotional distress)	2.0 (1.0)	3.2 (1.4)	-4.29	< 0.001
7 (family relationships)	3.3 (1.0)	3.6 (1.1)	-2.10	0.036
8 (autonomy)	3.0 (1.2)	3.1 (1.1)	-2.30	0.021
9 (frequency of leisure activities)	0.96 (0.63)	1.2 (0.74)	-2.12	0.034
10 (satisfaction with leisure activities)	3.4 (1.4)	3.9 (1.1)	-2.45	0.014

Data presented as mean (standard deviation).

GD = gambling disorder; G-SAS = Gambling Symptom Assessment Scale; SAS = Social Adjustment Scale; GFS-SR = Gambling Follow-up Scale, Self-Report version.

* Comprised the GFS-SR items gambling frequency, time spent gambling, money spent on gambling, and gambling craving, designated as the gambling behavior factor.

[†] Comprised the GFS-SR items family relationships, frequency of leisure activities, and satisfaction with leisure activities, designated as the social life factor.

[‡] Comprised the GFS-SR items debts, emotional distress, and autonomy, designated as the personal hardship factor.

and the Social Adjustment Scale total score ($\rho = 0.277$, $p < 0.002$).

Of the 50 patients who completed treatment, 30 (60%) did not fulfill the criteria for GD and were classified as recovered, 5 (10%) were classified as having mild GD, while 10 (20%) and 5 (10%) were classified as having moderate and severe GD, respectively. At this point, the total score of the GFS-SR and all its factor sub-scores were associated with the DSM-5 categories (p-values ranging from < 0.001 to 0.002), except for factor 2 (social life) ($\chi^2_{(3)} = 5.593$, $p = 0.133$). Additionally, all items from the GFS-SR, except for items 8 (autonomy), 9 (frequency of leisure activities), and 10 (satisfaction with leisure activities), were significantly associated with the DSM-5 categories (p-values ranging from < 0.001 to 0.026). Post-hoc analysis showed that the categories recovered and severe were most differentiated from the other categories, while the categories mild and moderate could not be differentiated from each other with reference to the GFS-SR variables. The total number of DSM-5 GD criteria met correlated strongly with the G-SAS ($\rho = 0.646$), the TLFB-GD hours ($\rho = 0.714$) and money ($\rho = 0.691$) scores, the GFS-SR total score ($\rho = -0.673$), and the Social Adjustment Scale total score ($\rho = 0.471$, all p-values ranging from 0.001 or lower). Finally, the DSM-5 GD criteria correlated significantly with all individual items of the GFS-SR (p-values ranging from < 0.001 to 0.015), except for items 8, 9, and 10.

Sensitivity to change over time

The Wilcoxon tests revealed significant differences between the pre- and post-treatment values for the number of DSM-5 GD criteria met, as well as for the GFS-SR total

score, GFS-SR factor 1 score (sum of items 1 to 4), GFS-SR factor 2 score (sum of items 7, 9, and 10), GFS-SR factor 3 score (sum of items 5, 6, and 8), G-SAS total score, and Social Adjustment Scale total score. On item-by-item analyses, there were significant differences between pre- and post-treatment values for all of the GFS-SR items. Items 1 through 6, which are directly related to gambling behavior, showed greater improvement than did items 7 through 10, which are related to quality of life. Table 5 presents a summary of the main results of pre- versus post-treatment analyses. The mean values for all but one of the ASI-G items indicated significant improvement after treatment ($z = -4.71$ to -3.53 , $p < 0.001$ for all), the exception being the item days of worry, for which the difference approached significance ($z = -1.92$, $p = 0.055$). Mean values for all TLFB-GD items (days, time, and money spent on gambling) also indicated significant improvement ($z = -4.33$ to -3.72 , $p < 0.001$ for all).

Discriminant validity, outcome prediction, and clinical cutoff

Based on the number of DSM-5 GD criteria met at the end of treatment, we classified 30 patients as recovered and 20 as unrecovered. Of the 30 recovered patients, 10 (33%) were negative for all criteria, five (17%) were positive for one criterion, 10 (33%) were positive for two, and five (17%) were positive for three criteria. The recovered patients showed significantly higher GFS-SR total scores ($U = 141.0$, $p = 0.012$), as well as significantly higher scores for GFS-SR factor 1 ($U = 154.0$, $p = 0.025$), GFS-SR factor 3 ($U = 91.0$, $p < 0.001$), G-SAS ($U = 145.0$, $p = 0.015$), and Social Adjustment Scale ($U = 141.0$, $p = 0.016$). Differences between the recovered and

unrecovered patients were not significant for the GFS-SR factor 2 score, for any ASI-G item score, or for any TLFB-GD item score. In the final logistic regression model with variation scores (Nagelkerke $R^2 = 0.383$, $\chi^2 = 15.4$, $p < 0.001$), only GFS-SR factor 3 was identified as a predictive factor (odds ratio = 1.70, $p = 0.003$), correctly classifying 80.0% and 70.6% of recovered and unrecovered patients, respectively.

Of the patients in our sample, 36 (30%) had attended GA meetings on a regular basis (two or more meetings per month) at baseline ($n=120$). At post-treatment ($n=50$), this number dropped to 13 (26%), although the difference was not significant (Wilcoxon $Z = -1.284$, $p = 0.198$). In addition, there was no statistical difference for GA attendance between recovered and unrecovered patients before treatment or after treatment, although there was a post-treatment trend for recovered patients to have attended GA more frequently than did unrecovered patients (pre-treatment: $U = 258.0$, $p = 0.358$; post-treatment: $U = 226.5$, $p = 0.078$). At baseline, there had been no differences between the recovered and unrecovered patients regarding demographics or gambling-related variables, except GFS-SR factor 2, on which the recovered patients had scored higher than had the unrecovered patients ($U = 178.5$, $p = 0.016$).

On ROC analysis, the area under the curve was 0.826 (standard error [SE] = 0.074, $p < 0.001$). The best GFS-SR cutoff score was found to be between 32.7 and 33.6 (i.e., GFS-SR total score > 33), which showed a sensitivity and specificity of 87% and 80%, respectively, for recovery from GD.

Discussion

The GFS-SR showed excellent convergent validity with the instruments (scales and interviews) most widely studied and applied within the field of gambling, as well as with the Social Adjustment Scale, the gold standard for assessing social adjustment. Because the GFS-SR, G-SAS, and Social Adjustment Scale share content, addressing mainly aspects of gambling severity, it was expected that those three instruments would correlate well among themselves. However, comparison of the GFS-SR with the DSM-5 criteria scoring yielded interesting results. The associations between them were only marginal when treatment began, but robust at the end of treatment. One factor that may have contributed is that, prior to treatment, most individuals were understandably compressed between the moderate and severe DSM-5 GD categories, which may have worked as a ceiling effect, shadowing some relationships that were more easily observed once treatment was completed. Caution must be exercised in approaching the associations between DSM-5 GD severity categories and GFS-SR scores, because the number of individuals in some categories at the end of treatment was too small. Nonetheless, correlation analysis between number of DSM-5 GD criteria and GFS-SR scores was not affected by this problem, and further corroborates the ability of the GFS-SR to gauge gambling severity along a continuum. Moreover, the GFS-SR presented stronger correlations with social

adjustment scales than did the number of DSM-5 criteria, suggesting that the former is better suited for the assessment of GD severity. This finding underscores the shortcomings inherent to using pre- and post-treatment diagnostic criteria counting to assess treatment effects, as well as the need for specific instruments to measure gambling severity and treatment efficacy. Thus, compared to its closest competitors, the GFS-SR shows some interesting peculiarities. Alongside the G-SAS, it is one of the few scales to address changes over the course of treatment in a self-report format. However, the GFS-SR seems to have a more balanced distribution of explained variance and items throughout its factor structure. Moreover, factor 2 (social life) encompasses aspects from the social environment (leisure activities and family relationships), which are not addressed by any other gambling-specific scale, including the DSM-5, and which have been shown to have a significant relationship with gambling recovery.

The GFS-SR also proved to be a reliable instrument, given that significant agreement was observed between collateral informant reports and patient self-reports. However, this finding must be viewed in light of the fact that patients were allowed to choose their collateral informant, and thus might have chosen a relative who would be likely to agree with their reports. In this aspect, there are ethical constraints, because selecting a relative without the consent of the patient would have constituted a breach of confidentiality. Unlike substance use, gambling cannot be detected by laboratory screening. Therefore, collateral informant reporting is still the best way to secure some insight into the consistency of self-reporting by gamblers.²¹ One piece of evidence against the assumption that gamblers and their relatives are prone to agree with each other is the fact that, in the present study, the GFS-SR question presenting the lowest kappa value was that concerning family relationships (item 7). Item 7 also showed the lowest convergent validity. Family relationships are probably too complex to be evaluated by a single question. We observed a tendency for pathological gamblers to view their families as less supportive than the relatives believed themselves to be. This finding probably reflects difficulties in the interaction between gamblers and their families, and underscores the need for family-oriented interventions as a regular component of the usual gambling treatment strategy.³² Unfortunately, additional insight into inter-rater agreement is limited by the fact that patients and collateral informants used correlated but different instruments. We could not employ a test-retest design, because the time interval would have differed between the patients and collateral informants. In addition, the psychometric properties of the collateral informant scale remain to be ascertained. Therefore, score reliability testing for gambling self-report measures relies mainly on analysis of internal consistency (Cronbach's alpha).

The structure of the GFS-SR is sound, with all 10 items contributing to the internal consistency of the scale. The three-factor structure revealed in the factor analysis mirrors the main goals of the treatment of behavioral syndromes^{2,33}: reduction of target symptoms (corresponding to GFS-SR factor 1); improvement of social adjustment (corresponding to GFS-SR factor 2); and reduction of subjective distress

(corresponding to GFS-SR factor 3). The three factors were comparable in terms of the proportion of variance explained, suggesting similar contributions to the clinical assessment of gamblers undergoing treatment.

The GFS-SR total score and factor scores, as well as the scores for every GFS-SR item, showed sensitivity to change. Over the 6-month treatment period, the GFS-SR items closely related to gambling behavior showed the greatest improvement. It is likely that improvements in the GFS-SR items related to family, autonomy, and leisure activities would be best observed during long-term, post-treatment follow-up.

In addition to its sensitivity to change, the GFS-SR revealed significant differences between recovered and unrecovered patients in terms of score variation over the course of treatment, thus precisely identifying the recovered patients, as did the G-SAS and the Social Adjustment Scale. The only exception was the score for GFS-SR factor 2, which, as previously stated, probably needs time beyond treatment to display consistent improvement. Nevertheless, a better baseline score for GFS-SR factor 2 was associated with less GD severity at the start of treatment and higher likelihood of recovery at treatment completion, suggesting that better social involvement is predictive of treatment response. Pre-treatment measures that can predict treatment response are rarely reported in the treatment of addictions. In GD, impulsivity and comorbid abuse of alcohol or drugs have been found to correlate negatively with treatment outcome.³⁴ Better social support has been shown to correlate significantly with the maintenance of abstinence in subjects attending GA meetings.³⁵ A protective effect against distressing experiences and gambling relapses has been attributed to good quality of life at post-treatment.³⁶ To our knowledge, ours is the first report of social support prior to treatment being associated with a positive outcome, which further underscores the relevance of the social environment for recovery from GD. In addition, we found that a decrease in gambling-related distress (i.e., an increase in the GFS-SR factor 3 score) was the best predictor of gambling recovery, with a one-point increase in the GFS-SR factor 3 score increasing the chance of recovery by 70%. This finding provides, for the first time, empirical support for the Banff Consensus assertion that addressing the negative consequences of gambling is as important for clinical recovery as is curtailing the gambling behavior itself.³ Furthermore, the almost-perfect convergence between the outcomes presented herein, based on the DSM-5 criteria, and those observed in the alternative analysis, based on the DSM-IV-TR criteria and diagnostic threshold, is an indication that researchers and clinicians referring to either classification manual can use the GFS-SR cutoff score of 33 to identify gambling recovery.

The difference between recovered and unrecovered patients in terms of GA attendance showed only borderline significance. However, the small size of our sample might have curtailed a potential synergistic interaction between clinical and self-help approaches. Previous studies have suggested that the combination of GA and clinical treatment increases gambling abstinence rates.³⁷ In addition, the validation study of the first version of the

GFS showed that GA participation increased leisure frequency among gamblers.²³ Compared to its first version, the GFS-SR retained its psychometric robustness, with the advantage of switching from a clinician-rated to a self-report format. The original GFS had a single-factor structure, whereas the GFS-SR presents a three-factor structure, stemming from its wider variety of items that better mirror the underlying dimensions of gambling recovery. Despite its greater complexity, the GFS-SR remained as easy to use as the original scale. It is also noteworthy that, in the first validation study, frequency of leisure correlated significantly with gambling abstinence. In the current study, factor 2 (social life), which comprises three items – two of them addressing leisure – was the sole predictor of gambling recovery, which is in keeping with findings from our own group about the relevance of quality of life for the retention of treatment gains.^{38,39}

Our study has certain limitations. First, despite the fact that the translation of the DSM-IV-TR criteria used in this study has been extensively used in previous clinical and epidemiological studies conducted by our group,⁴⁰ this Portuguese version still lacks formal validation. Second, the small size of the patient sample and the convenience strategy employed limited the statistical power of the study. Third, assessment of family relationships by the GFS-SR is less than ideal. It is likely that the scale could be improved by the addition of a few more questions addressing this topic, which would require further validation studies. Fourth, a considerable proportion of eligible subjects either declined or dropped out of treatment; however, the fact that dropouts and treatment completers did not differ significantly at baseline partly compensated for this. Finally, the data presented in this study are related to the original Portuguese-language version of the GFS-SR (Appendices 1 and 2, available as online-only supplementary materials, represent free translations into English by the authors). Therefore, further studies are needed to check whether translations of the GFS-SR into English or other languages will retain the same psychometric properties.

The GFS-SR has moderate to excellent convergent validity with other scales that measure gambling. Its psychometrically sound, three-factor structure provides a measure that is mostly sensitive to changes in gambling behavior that occur during treatment (factor 1), significant social predictors of treatment response (factor 2), and a dependable measure of gambling distress that helps differentiate between treatment responders and non-responders (factor 3). In its current format, the GFS-SR works as a hybrid scale, assessing gambling content as well as social adjustment issues affected by GD. Although both of those topics are covered separately by the G-SAS and the Social Adjustment Scale, respectively, those scales take longer to apply than does the GFS-SR. In comparison with the TLFB-GD and ASI-G, the GFS-SR offers the alternative of a self-report assessment, thus circumventing interview bias on the report of sensitive topics (i.e., gambling relapses, debts, and emotional distress). Consequently, the GFS-SR seems to be a reliable instrument for the follow-up of and assessment of

treatment outcomes in GD patients, both in research and in clinical settings. Future studies should investigate the potential utility of the GFS-SR in other contexts related to gambler support, such as self-help and other lay interventions.

Acknowledgements

The authors are grateful to Cecilia Galetti, Fabiano Blasquez, and Juliana Ono Tonaki for their assistance with data input, as well as to the staff of the Gambling Outpatient Unit of the Institute and Department of Psychiatry, Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo. This study received financial support from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES; grant to AMG).

Disclosure

HT has worked as the chief gambling consultant for the Brazilian federal bank (Caixa Econômica Federal) since March 2015, overseeing its national lottery products, and as a member of the Independent Assessment Panel for the responsible gambling certification of the World Lottery Association from October 2015 to February 2016. The other author reports no conflicts of interest.

References

- Nowak DE, Aloe AM. The prevalence of pathological gambling among college students: a meta-analytic synthesis, 2005-2013. *J Gambli Stud.* 2014;30:819-43.
- Kushnir V, Cunningham JA, Hodgins DC. A prospective natural history study of quitting or reducing gambling with or without treatment: protocol. *JMIR Res Protoc.* 2013;2:e51.
- Walker M, Toneatto T, Potenza MN, Petry N, Ladouceur R, Hodgins DC, et al. A framework for reporting outcomes in problem gambling treatment research: the Banff, Alberta Consensus. *Addiction.* 2006;101:504-11.
- Grant JE, Odlaug BL, Schreiber LR. Pharmacological treatments in pathological gambling. *Br J Clin Pharmacol.* 2014;77:375-81.
- Cowlshaw S, Merkouris S, Dowling N, Anderson C, Jackson A, Thomas S. Psychological therapies for pathological and problem gambling. *Cochrane Database Syst Rev.* 2012;11:CD008937.
- Hodgins DC, Currie S, el-Guebaly N, Peden N. Brief motivational treatment for problem gambling: a 24-month follow-up. *Psychol Addict Behav.* 2004;18:293-6.
- Ladouceur R, Sylvain C, Boutin C, Lachance S, Doucet C, Leblond J. Group therapy for pathological gamblers: a cognitive approach. *Behav Res Ther.* 2003;41:587-96.
- Melville CL, Davis CS, Matzenbacher DL, Clayborne J. Node-link-mapping-enhanced group treatment for pathological gambling. *Addict Behav.* 2004;29:73-87.
- Petry NM, Weinstock J, Ledgerwood DM, Morasco B. A randomized trial of brief interventions for problem and pathological gamblers. *J Consult Clin Psychol.* 2008;76:318-28.
- Lesieur HR, Blume SB. The South Oaks Gambling Screen (SOGS): a new instrument for the identification of pathological gamblers. *Am J Psychiatry.* 1987;144:1184-8.
- Stinchfield R. Reliability, validity, and classification accuracy of the South Oaks Gambling Screen (SOGS). *Addict Behav.* 2002;27:1-19.
- Hodgins DC. Using the NORC DSM Screen for Gambling Problems as an outcome measure for pathological gambling: psychometric evaluation. *Addict Behav.* 2004;29:1685-90.
- Kim SW, Grant JE, Potenza MN, Blanco C, Hollander E. The Gambling Symptom Assessment Scale (G-SAS): a reliability and validity study. *Psychiatry Res.* 2009;166:76-84.
- American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR)* Arlington: American Psychiatric Publishing; 2000.
- American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5)* Arlington: American Psychiatric Publishing; 2013.
- Romanczuk-Seiferth N, van den Brink W, Goudriaan AE. From symptoms to neurobiology: pathological gambling in the light of the new classification in DSM-5. *Neuropsychobiology.* 2014;70:95-102.
- McLellan AT, Kushner H, Metzger D, Peters R, Smith I, Grissom G, et al. The fifth edition of the Addiction Severity Index. *J Subst Abuse Treat.* 1992;9:199-213.
- Petry NM. Validity of a gambling scale for the addiction severity index. *J Nerv Ment Dis.* 2003;191:399-407.
- Petry NM. Concurrent and predictive validity of the Addiction Severity Index in pathological gamblers. *Am J Addict.* 2007;16:272-82.
- Sobell LC, Sobell MB. Timeline follow-back: a technique for assessing self-report alcohol consumption. In: Litten R, Allen J, editors. *Measuring alcohol consumption: psychosocial and biochemical methods.* Totowa: Humana Press; 1992. p. 41-72.
- Hodgins DC, Makarchuk K. Trusting problem gamblers: reliability and validity of self-reported gambling behavior. *Psychol Addict Behav.* 2003;17:244-8.
- Weinstock J, Whelan JP, Meyers AW. Behavioral assessment of gambling: an application of the timeline follow-back method. *Psychol Assess.* 2004;16:72-80.
- de Castro V, Fuentes D, Tavares H. The gambling follow-up scale: development and reliability testing of a scale for pathological gamblers under treatment. *Can J Psychiatry.* 2005;50:81-6.
- World Medical Association. *World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects.* *J Postgrad Med.* 2002;48:206-8.
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33:159-174.
- Weissman MM, Bothwell S. Assessment of social adjustment by patient self-report. *Arch Gen Psychiatry.* 1976;33:1111-5.
- Gorenstein C, Moreno RA, Bernik MA, Carvalho SC, Nicastrí S, Cordás T, et al. Validation of the Portuguese version of the Social Adjustment Scale on Brazilian samples. *J Affect Disord.* 2002;69:167-75.
- Flaherty JA, Gaviria FM, Pathak D, Mithell T, Wintrob R, Richman JA, et al. Developing instruments for cross-cultural psychiatric research. *J Nerv Ment Dis.* 1988;176:257-63.
- Tavares H, Gentil V, Oliveira CS, Tavares AG. Jogadores patológicos, uma revisão: psicopatologia, quadro clínico e tratamento. *Rev Psiquiatr Clin.* 1999;26:179-87.
- Ladouceur R, Sylvain C, Boutin C, Lachance S, Doucet C, Leblond J, et al. Cognitive treatment of pathological gambling. *J Nerv Ment Dis.* 2001;189:774-80.
- Statistical Package for Social Sciences (SPSS). *Statistical package for social sciences.* Version 10.0 for Windows Chicago: SPSS Inc; 1997.
- Lee BK, Awosoga O. Congruence couple therapy for pathological gambling: a pilot randomized controlled trial. *J Gambli Stud.* 2015;31:1047-68.
- Grayson SN. Role of outcome measurement in psychiatry. In: In: Hak WW, Burt T, Sederer LI, editors. *Outcome measurement in psychiatry: a critical review.* New York: American Psychiatric Publishing; 2002. p. 3-13.
- Melville KM, Casey LM, Kavanagh DJ. Psychological treatment dropout among pathological gamblers. *Clin Psychol Rev.* 2007;27:944-58.
- Oei TP, Gordon LM. Psychosocial factors related to gambling abstinence and relapse in members of gamblers anonymous. *J Gambli Stud.* 2008;24:91-105.
- Sander W, Peters A. Pathological gambling: influence of quality of life and psychological distress on abstinence after cognitive-behavioral inpatient treatment. *J Gambli Stud.* 2009;25:253-62.
- Petry NM, Ammerman Y, Bohl J, Doersch A, Gay H, Kadden R, et al. Cognitive-behavioral therapy for pathological gamblers. *J Consult Clin Psychol.* 2006;74:555-67.
- Magalhães AC, Jungerman FS, Silva MC, Moraes MM, Tavares H. Post-therapy group for pathological gamblers: improvement beyond symptoms. *Rev Bras Psiquiatr.* 2009;31:181-2.
- Angelo DL, Tavares H, Zilberman ML. Evaluation of a physical activity program for pathological gamblers in treatment. *J Gambli Stud.* 2013;29:589-99.
- Tavares H. Gambling in Brazil: a call for an open debate. *Addiction.* 2014;109:1972-6.