Psychological Assessment

Initial Development and Psychometric Properties of the Gambling Disorder Test in a Nationally Representative Sample of Adults

Halley M. Pontes, Špela Selak, Mark Žmavc, and Mark D. Griffiths

Online First Publication, March 24, 2025. https://dx.doi.org/10.1037/pas0001374

Psychological Assessment

ISSN: 1040-3590
https://doi.org/10.1037/pas0001374

Initial Development and Psychometric Properties of the Gambling Disorder Test in a Nationally Representative Sample of Adults

Halley M. Pontes[1], Špela Selak[2], Mark Žmavc[3], and Mark D. Griffiths[4]
[1] School of Psychological Sciences, Birkbeck, University of London
[2] National Institute of Public Health, Ljubljana, Slovenia
[3] Centre for Digital Wellbeing Logout, Ljubljana, Slovenia
[4] International Gaming Research Unit, Psychology Department, Nottingham Trent University, Nottingham, United Kingdom

Gambling disorder (GD) is an officially recognized mental health disorder. However, its conceptualization and diagnostic criteria have changed substantially over the years due to new clinical and epidemiological research supporting its reconceptualization from an impulse control disorder to an addictive disorder. The evolving nature of GD led to changes in its diagnostic approach within the 11th revision of the *International Statistical Classification of Diseases and Related Health Problems* (*ICD-11*). However, no updated standardized psychometric test reflecting the latest developments exists. Therefore, the goal of the present study was to develop and report the psychometric properties of the Gambling Disorder Test (GDT), a brief and convenient four-item assessment instrument reflecting the current diagnostic criteria for GD in the *ICD-11*. A nationally representative sample of British adults was recruited (N = 1,028, M_{age} = 46.54 years, SD_{age} = 15.71). The results showed a one-factor solution for the GDT and initial support for the scale's factorial validity, population cross-validity, criterion validity, concurrent validity, and reliability. Further gender-based measurement invariance was conducted, with the GDT exhibiting full scalar invariance and the results of latent mean comparison showing that males had significantly higher GD latent means compared to females (latent mean difference = −0.156; $z = -3.844$, $p < .001$, $d = -.249$). The self-reported prevalence of GD in the sample was 0.49%. The GDT is a promising brief assessment instrument based on the latest conceptualization and diagnostic criteria for GD that can be employed by clinicians and researchers alike.

Public Significance Statement
The newly developed Gambling Disorder Test, based on the World Health Organization conceptualization for gambling disorder offers a brief, valid, and reliable psychometric instrument for assessing gambling disorder based on the latest conceptualization and diagnostic criteria. The Gambling Disorder Test can significantly aid clinicians and researchers in identifying and understanding gambling disorder, facilitating better treatment and research outcomes.

Keywords: Gambling Disorder, Problem Gambling, Gambling, World Health Organization, 11th revision of the *International Statistical Classification of Diseases and Related Health Problems*

Supplemental materials: https://doi.org/10.1037/pas0001374.supp

Ryan J. Marek served as action editor.
Halley M. Pontes https://orcid.org/0000-0001-8020-7623
The authors declare no conflicts of interest. This work has been funded by an internal grant from Birkbeck, University of London, that has been granted to the corresponding author, Halley M. Pontes. Mark D. Griffiths has received research funding from Norsk Tipping (the gambling operator owned by the Norwegian government). Mark D. Griffiths has received funding for a number of research projects in the area of gambling education for young people, social responsibility in gambling, and gambling treatment from Gamble Aware (formerly the Responsibility in Gambling Trust), a charitable body which funds its research program based on donations from the gambling industry. Mark D. Griffiths undertakes consultancy for various gambling companies in the area of player protection and social responsibility in gambling.

Halley M. Pontes played a lead role in conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, software, supervision, validation, visualization, writing–original draft, and writing–review and editing. Špela Selak played a supporting role in project administration, writing–original draft, and writing–review and editing and an equal role in conceptualization, investigation, supervision, and validation. Mark Žmavc played a supporting role in conceptualization, investigation, project administration, supervision, writing–original draft, and writing–review and editing. Mark D. Griffiths played a supporting role in conceptualization, investigation, methodology, project administration, writing–original draft, and writing–review and editing and an equal role in supervision.
Correspondence concerning this article should be addressed to Halley M. Pontes, School of Psychological Sciences, Birkbeck, University of London, Malet Street, Bloomsbury, London WC1E 7HX, United Kingdom. Email: contactme@halleypontes.com

Gambling is a popular recreational activity that allows individuals to place bets or stakes, typically involving money or items of monetary value, with the potential to gain additional financial rewards (Sirola et al., 2023). Popular gambling activities include lotteries, scratchcards, sports betting, bingo, casino games, and slot machines. Although gambling is a harmless activity for most people, a minority of individuals develop problematic patterns of gambling behavior leading to the experience of emotional, social, and financial harms affecting not only them but also their significant others (Langham et al., 2015). Given its hazards, jurisdictions around the world are becoming increasingly interested in protecting their citizens from gambling harms by regulating the activity and increasing public awareness in relation to problem gambling (PG). For instance, in 2017, the Gambling Commission in the United Kingdom classed PG as a public health concern (Wardle et al., 2019).

In terms of the prevalence rates, a systematic review reported that prevalence rates for PG between 2000 and 2015 varied significantly across different countries worldwide (range: 0.12%–5.8%) while in Europe, the prevalence ranged from 0.12% to 3.40% (Calado & Griffiths, 2016). A more recent meta-analysis by Gabellini et al. (2023) reviewing the evidence from studies published between 2016 and 2022 reported a prevalence rate for PG of 1.19%.

In regard to its conceptualization, the definition of PG has evolved significantly over the past decades, especially when considering existing diagnostic manuals such as the American Psychiatric Association *Diagnostic and Statistical Manual of Mental Disorders* (*DSM*) and the World Health Organization (WHO) *International Statistical Classification of Diseases and Related Health Problems* (*ICD*). Historically, PG has been included in the *ICD* since 1975 and in the *DSM* since 1980 (Abbott, 2020). Within both the *DSM* and *ICD*, PG has been described using different terminology and has undergone numerous changes to its core diagnostic criteria and definition. For instance, the American Psychiatric Association (2013) reclassified PG in the fifth revision of the *DSM* (*DSM-5*) from an impulse control disorder to an addictive disorder due to its shared commonalities with different aspects of substance use disorders (Reilly & Smith, 2013).

Within the American Psychiatric Association diagnostic framework, the current *Diagnostic and Statistical Manual of Mental Disorders*, fifth edition, text revision (American Psychiatric Association, 2022) defines gambling disorder (GD) as a persistent and recurrent problematic gambling behavior, resulting in significant impairment or distress that can be diagnosed when individuals exhibit at least four of the following symptoms within a 12-month period: (a) the need to gamble increasing amounts of money to achieve excitement, (b) restlessness or irritability when trying to reduce or stop gambling, (c) repeated unsuccessful efforts to control gambling, (d) preoccupation with gambling, (e) gambling to escape distress, (f) returning to gambling after losing money to recover losses, (g) lying to conceal gambling involvement, (h) jeopardizing significant relationships or opportunities due to gambling, and (i) relying on others for financial relief from gambling-related problems. The severity of these symptoms can vary depending on the number of criteria endorsed, ranging from mild (4–5 criteria), moderate (6–7 criteria), to severe (8–9 criteria).

Similarly, the WHO diagnostic framework defines GD in the *ICD-11* (World Health Organization, 2022b) as a pattern of persistent or recurrent gambling behavior over a 12-month period that can be diagnosed when individuals endorse the following diagnostic criteria: (a) impaired control over gambling (e.g., onset, frequency, intensity, duration, termination, context), (b) increasing priority given to gambling to the extent that it takes precedence over other life interests and daily activities, (c) continuation or escalation of gambling despite the occurrence of negative consequences, and (d) experience of significant distress and/or impairment in personal, family, social, educational, occupational, or other important areas of functioning due to the gambling behavior (World Health Organization, 2022a).

At present, several long- and short-form screening instruments exist for assessing the severity of GD. However, existing GD measures present several important diagnostic and theoretical shortcomings. In a recent systematic review of 31 screening tools, Otto et al. (2020) reported that only three instruments had been validated against an adequate reference standard. Notably, these instruments only included items based on older diagnostic criteria such as those from the *DSM-IV* and *DSM-5*. Consequently, Otto et al. (2020) concluded that at present, very few instruments for GD have been validated with sufficient methodological quality to be recommended for use across the health system. Similarly, Dowling et al. (2019) conducted a meta-analysis of brief screening instruments for GD and identified a total of 20 instruments, with 10 showing sufficient evidence to detect GD and only five being suitable for detecting both problem and at-risk gambling.

Existing popular screening instruments that are widely used in the field, such as the South Oaks Gambling Screen (SOGS; Lesieur & Blume, 1987) and the Problem Gambling Severity Index (PGSI; Ferris & Wynne, 2001), also present significant limitations. More specifically, the SOGS is a 20-item instrument that was developed based on the *DSM-III* criteria for PG in 1980 and has been reported to produce high rates of false-positive diagnosis, overestimating the true number of affected individuals in the general population (Stinchfield, 2002). Moreover, the PGSI comprises a set of nine items developed based on the *DSM-IV* criteria for PG, published in 1994 combined with items similar to those in the SOGS, making it arguably the most frequently used screening instrument for PG worldwide (Caler et al., 2016) despite its reliance on outdated diagnostic criteria.

Given the aforementioned rationale, it is evident that there is a need for a brief screening instrument for GD based on the latest advancements in the field, such as the WHO diagnostic framework for GD proposed in the *ICD-11* (World Health Organization, 2022a). The development of such an instrument builds upon and enhances the validity of existing GD assessments by integrating the latest diagnostic criteria outlined in the *ICD-11* and aligning the assessment of GD with the current view that GD is best conceptualized as an addictive disorder as reported in previous studies (Kim & Hodgins, 2019; Saunders et al., 2017). Therefore, the present study sought to conduct an initial development of a concise and user-friendly approach to assessing GD that captures the full scope of GD's current diagnostic framework positioning the construct as an addictive disorder. Notably, the development of a new assessment instrument for GD based on the recent *ICD-11* criteria has the potential to enable more accurate differentiation of disordered gambling levels among both community-based and clinical samples while also providing clinicians and researchers with a robust, updated psychometric test that captures relevant GD symptoms with greater precision and specificity, thereby expanding the range of validated instruments available to meet diverse clinical and research needs.

To the best of the authors' knowledge, only one current psychometric instrument exists for GD based on the WHO diagnostic framework. The Assessment of Criteria for Specific Internet-use Disorders (ACSID-11; Müller et al., 2022) includes 11 items based on the *ICD-11* criteria for disorders due to addictive behaviors. However, this measure was not developed exclusively for GD as it assesses five online potentially addictive behaviors and focuses solely on internet-use disorders. As such, the ACSID-11 is arguably not a parsimonious instrument and it may not be optimal for assessing land-based gambling behaviors. Therefore, the present study addresses this gap by reporting the initial development and psychometric properties of a dedicated, *ICD-11*-based assessment tool specifically designed for GD that offers a concise yet robust assessment, directly supporting clinicians and researchers in identifying and understanding gambling-related issues within the most up-to-date diagnostic framework.

Building on these insights, the main goal of the present study was to develop a brief, standardized psychometric instrument for GD irrespective of where the gambling activity is carried out (e.g., land-based or online), grounded in the latest clinical and diagnostic advancements and aligned with the WHO framework outlined in the *ICD-11*. This updated approach is expected to yield significant benefits for the assessment of GD in both clinical practice and research contexts.

Method

Participants and Procedure

The present study recruited a large and nationally representative sample of 1,074 individuals from the United Kingdom via *Prolific* (https://www.prolific.com) through an online survey developed in *Qualtrics*. The recruitment took place between May 17 and 19, 2024, with each participant receiving a financial compensation (average reward per hour: £9.21/hr) for their participation. The study obtained ethical approval from the Faculty of Science Ethics Committee at Birkbeck, University of London (Reference No. 2324063) and was conducted according to the guidelines of the Declaration of Helsinki.

To enhance the quality of the data collected, two types of controls were implemented to ensure the validity of the data. First, two attention checks were added at different intervals in the online survey around the middle (50^{th} percentile) and the end (i.e., 75^{th} percentile) to mitigate the potential impact of inattention. These two attention checks involved asking participants to answer a question with *agree* and another one with *strongly agree*. Second, participants who completed the survey exceptionally fast (i.e., three standard deviations below the mean completion time) were excluded to further improve data quality. As a result of this data quality control procedure, a total of 46 (4.28%) cases were excluded, leading to a final sample size of 1,028 participants ($M_{age} = 46.54$ years, $SD_{age} = 15.71$, 50.97% female) that were used for the present analyses.

The process utilized by *Prolific's* representative sampling tool ensures that the recruited sample reflects the national demographic distribution of the population based on age, sex, and ethnicity. To achieve this, census data from the U.K. Office of National Statistics is used to stratify the sample so that it matches the proportions of these demographics in the general population. The advantage of recruiting a nationally representative sample is that the study results provide a robust basis for making inferences about the U.K. population.

Measures

Sociodemographic and Gambling-Related Variables

Basic sociodemographic data were collected about participants' characteristics and gambling-related behaviors (see Table 1). These included age, gender, educational level, employment status, relationship status, gambling activity in the last 12 months, average amount of money spent gambling in the past 12 months, average hours per week and weekend spent gambling, perceived level of engagement with gambling, complaints from others due to gambling behaviors, and experience of a significant problem in life in the last 12 months due to gambling.

Table 1
Participants' Sociodemographic Characteristics and Gambling-Related Behaviors

Variable	Statistic
Age (years), *M* (*SD*)	46.54 (15.71)
Gender, *n* (%)	
Female	524 (50.97)
Male	495 (48.15)
Other	9 (0.88)
Education level, *n* (%)	
No formal qualifications	14 (1.36)
GCSEs or equivalent (e.g., O-levels)	138 (13.42)
A-levels or equivalent (e.g., BTEC)	176 (17.12)
Vocational qualification (e.g., NVQ, HND, HNC)	121 (11.77)
Undergraduate degree (e.g., BA, BSc)	388 (37.74)
Postgraduate degree (e.g., MA, MSc, PhD, PGCE)	191 (18.58)
Employment status, *n* (%)	
Unemployed	278 (27.04)
Employed full-time	512 (49.81)
Employed part-time	238 (23.15)
Relationship status, *n* (%)	
Single	266 (25.88)
In a relationship	244 (23.74)
Married	438 (42.61)
Separated	18 (1.75)
Divorced	44 (4.28)
Widowed	18 (1.75)
Gambled in the past 12 months? *n* (%)	
No	530 (51.56)
Yes	498 (48.44)
Average monthly amount spent gambling past 12 months, *M* (*SD*)[a]	197.87 (806.03)
Average hours per week spent gambling, *M* (*SD*)[a]	2.31 (5.93)
Average hours per weekend spent gambling, *M* (*SD*)[a]	2.24 (9.91)
Perceived level of engagement with gambling, *n* (%)[a,b]	
Casual	362 (72.69)
Regular	96 (19.28)
Frequent	33 (6.63)
Intense	7 (1.41)
Complaints from others due to gambling behaviors, *n* (%)[a,b]	
No	468 (93.98)
Yes	30 (6.02)
Significant problem experienced in the last 12 months due to gambling, *n* (%)[a,b]	
No	474 (95.18)
Yes	24 (4.82)
Gambling disorder prevalence with the GDT, *n* (%)	5 (0.49)

Note. Some values may not total up to 100 due to rounding. GCSE = General Certificate of Secondary Education; GDT = Gambling Disorder Test.
[a] Data concerning only those who reported having gambled in the past 12 months ($n = 498$). [b] Percentages provided only in relation to those who reported having gambled in the past 12 months.

The Gambling Disorder Test

The Gambling Disorder Test (GDT) is a brief standardized psychometric instrument designed to assess GD (Code: 6C50) according to the WHO diagnostic framework proposed in the *ICD-11* (World Health Organization, 2022b). The overall development process for the GDT was rooted in a direct adaptation of the *ICD-11* diagnostic criteria for GD. Rather than developing a broad bank of potential items, each *ICD-11* criterion was carefully examined and rephrased into a clear and concise item suitable for a brief psychometric instrument. This approach ensured that the final items were directly aligned with the four core diagnostic features of GD as defined by the WHO, emphasizing clinical relevance, as well as diagnostic, and conceptual alignment. By closely aligning to the *ICD-11* criteria, the present study aimed to develop an instrument that remains succinct and relevant to both clinicians and researchers interested in the assessment of GD.

As a result of this process, the GDT includes four items answered on a 5-point Likert scale (1 = *never*; 2 = *rarely*; 3 = *sometimes*; 4 = *often*; and 5 = *very often*). Total scores are calculated by summation of participant's answers given to all items and can range from 4 to 20, with greater scores reflecting higher levels of GD symptoms. For research purposes, answering all four items with 4: *often* or 5: *very often* can be operationalized as endorsement of a specific diagnostic criterion. Therefore, by adopting this approach, researchers may be able to identify potentially disordered and nondisordered gamblers within their samples.

In order to develop the items of the GDT, the psychometric test development guidelines recommended by the Standards for Educational and Psychological Testing (American Educational Research Association, 2014) were followed. An initial set of items that mapped directly onto the *ICD-11* GD criteria (World Health Organization, 2022a) was devised and sent to a small group of experts ($n = 5$) with a track record of published research on GD to provide feedback about the wording of each item to help enhance the test's content validity, which is key to ensure that all items match the conceptual definition for GD and include the appropriate theoretical and practical considerations (Hair et al., 2010). Following this, a pilot survey was conducted to examine the face validity of the GDT items among a sample of 26 individuals (53% male, $M_{\text{age}} = 18.30$ years, $SD_{\text{age}} = 4.10$ years) who reported having gambled in the past 12 months, with no participant reporting significant issues when answering all new items.

Once feedback was collected from experts and nonexperts, minor revisions to the wording of the items were conducted in line with the feedback provided in order to improve the concept coverage and comprehension of the items. The final set of items of the GDT and their respective diagnostic criteria are the following (see Supplemental Table S1): Item 1: "I have had difficulties controlling my gambling activity" (impaired control); Item 2: "I have given increasing priority to gambling over other life interests and daily activities" (increased priority), Item 3: "I have continued gambling despite the occurrence of negative consequences" (continuation despite negative consequences), and Item 4: "I have experienced significant problems in life (e.g., personal, family, social, education, occupational) due to the severity of my gambling behavior" (functional impairments).

Problem Gambling Severity Index

The PGSI (Ferris & Wynne, 2001) is arguably one of the most widely utilized measures of PG. It includes a total of nine items asking participants about their gambling activities in the past 12 months. Each item is responded to on a 4-point Likert scale (0 = *never* to 3 = *almost always*) and total scores can range from 0 to 27, with higher scores suggesting greater levels of PG. Participants' risk for PG can be inferred from total scores based on the following score ranges: 0: nonproblem gambler, 1–2: low risk, 3–7: moderate risk, and ≥8: problem gambler (Ferris & Wynne, 2001).

Statistical Analyses and Analytic Strategy

The main statistical analyses comprised of (a) descriptive statistics of participants' demographic and gambling-related characteristics; (b) inter-item and cross-item correlational analyses between all items of the GDT and PGSI; and (c) psychometric analyses of the GDT's factorial validity and population cross-validity utilizing exploratory factor analysis (EFA) and confirmatory factor analysis (CFA), reliability analysis utilizing Cronbach's alpha (α), McDonald's Omega (Ω), and less used coefficients of internal consistency such as composite reliability (CR) and factor determinacy (FD), measurement invariance across genders, and further validity testing of the GDT by testing its associations with relevant gambling-related measures. The Supplemental Material reports the findings of item response theory (IRT) analysis using the graded response model (GRM) to further evaluate the psychometric performance of the GDT beyond the classical test theory (CTT) and across varying levels of GD severity, providing the relevant item information curves (IICs) and test information curve (TIC) results.

For the reliability analysis, the CR coefficient measures the internal consistency of a set of indicators taking into account factor loadings and measurement errors of the indicators within a measurement model while the FD coefficient measures how well a factor score represents the underlying latent variable. Typically, values of CR and FD closer to one indicate high levels of reliability. However, the adequacy of CR and FD coefficients was determined based on the commonly adopted thresholds of ≥.70 (Fornell & Larcker, 1981; Hair et al., 2010) and ≥.80, respectively (Muthén & Muthén, 2018).

For measurement invariance, configural, metric, and scalar invariance was tested using multigroup CFA across gender adopting standard psychometric procedures (Putnick & Bornstein, 2016). First, configural invariance (i.e., equal factor structure) was tested by estimating a baseline model to assess whether the same factor structure is equivalent between the groups. If configural invariance is supported, then metric invariance (i.e., constrained factor loadings) was tested by imposing equality constraints on factor loadings while allowing other parameters to be freely estimated. If metric invariance is achieved, then scalar invariance (i.e., equal factor loadings and intercepts) was tested. In case scalar invariance is demonstrated across the groups, then mean differences in GDT can be attributed to true differences in GD symptoms across genders. As such, if scalar invariance is supported, then latent mean comparison between both genders was conducted using the Wald test to determine the statistical significance of the difference between the latent means. To assess lack of measurement invariance, fit indices of the estimated nested models were compared based on changes observed in comparative fit index (CFI) $\geq -.010$ in addition to a change in root-mean-square error of approximation (RMSEA) of ≥.015 (Chen, 2007; Cheung & Rensvold, 2002).

Prior to the analyses, the full dataset was split into two age- and gender-matched subsamples in order to assess whether the factorial solution found in the EFA using one sample can be replicated with CFA in another sample to examine population cross-validity (Raju et al., 1997, 1999), an approach that has been previously employed in the field (see Pontes & Griffiths, 2015). The splitting of the Complete Sample was performed using the *matchit*() function which uses propensity score matching with the nearest neighbor matching method. As a result, the complete sample was split into Sample 1 (n = 514, M_{age} = 45.50 years, SD_{age} = 15.79) and Sample 2 (n = 514, M_{age} = 47.58 years, SD_{age} = 15.58, see Supplemental Table S2). The quality of the estimated structural equation models was assessed using several fit indices and their recommended thresholds: χ^2/df (1, 4); CFI; and Tucker–Lewis fit index (TLI) [.90, .95]; RMSEA [<.05]; and the standardized root-mean-square residual (SRMR) [<.08] (Bentler, 1990; Bentler & Bonnet, 1980; Hooper et al., 2008; Hu & Bentler, 1999).

All the analyses were conducted in R Version 4.4.1 "Race for Your Life" (R Core Team, 2024) using the following packages: psych v 2.4.3 (Revelle, 2024), dplyr v. 1.14 (Wickham et al., 2023), tidyr v. 1.3.1 (Wickham et al., 2024); Hmisc v. 5.1-3 (Harrell & Dupont, 2024), GPArotation 2024.3-1 (Bernaards & Jennrich, 2005), MatchIt v. 4.5.5 (Ho et al., 2011), lavaan v. 0.6-18 (Rosseel, 2012), GGally v. 2.2.1 (Schloerke et al., 2024), and ggplot2 v. 3.5.1 (Wickham, 2016).

Transparency and Openness

This study's design and its analysis were not preregistered. The data that support the findings of this study will be made available to interested readers from the corresponding author upon reasonable request.

Results

Sample Characteristics and Key Gambling-Related Behaviors

Basic sociodemographic data and gambling-related behaviors of the Complete Sample are shown in Table 1. The gender split was relatively even between female (50.97%, n = 524) and male (48.15%, n = 495) participants, with a small fraction of participants reporting "Other" as their gender (0.88%, n = 9). In terms of educational level, 37.74% (n = 388) of all participants reported having completed an undergraduate degree (e.g., BA, BSc). Nearly half of the sample (49.81%, n = 512) were employed full-time and 27.04% (n = 278) were unemployed.

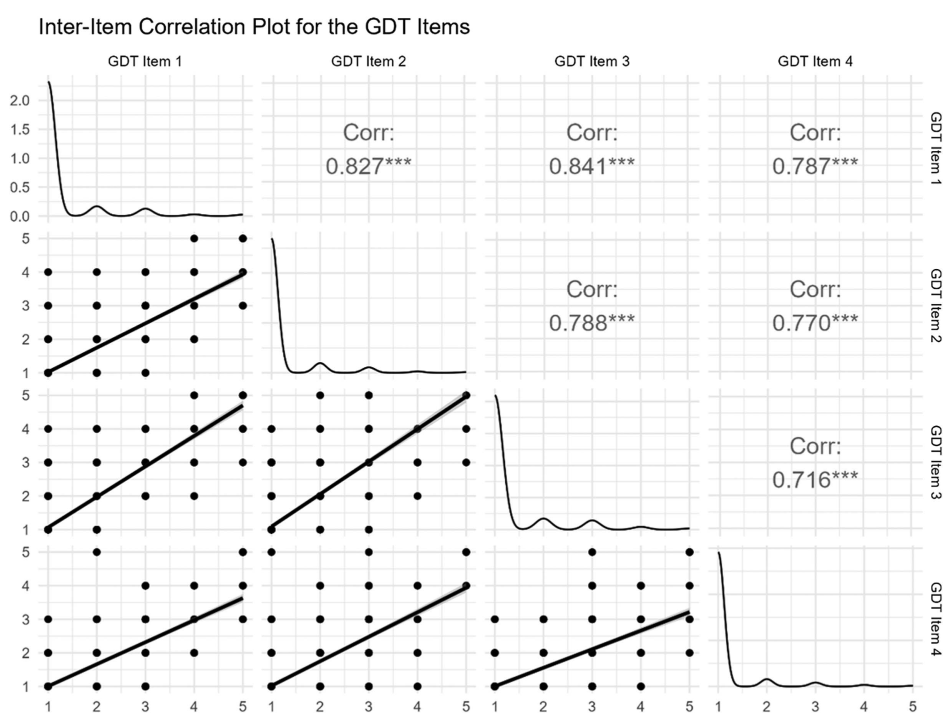
In terms of gambling-related behaviors, nearly half of the sample reported having gambled in the past 12 months (48.44%, n = 498), with 6.63% (n = 33) declaring being a "frequent gambler" and 1.41% (n = 7) reporting being an "intense gambler." Moreover, 6.02% (n = 30) of the total sample reported that others had complained about their gambling activity with a further 4.82% (n = 24) reporting they had experienced a significant problem in the last 12 months due to their gambling activity. Overall, the self-reported prevalence rate of GD assessed with the GDT was 0.49% (n = 5).

Inter-Item and Cross-Item Correlational Analyses

As can be seen in Figure 1, the inter-item correlations of the four items GDT items were strong and statistically significant at $p < .001$, supporting the instrument's internal consistency and coherence

Figure 1

Inter-Item Correlation Analysis All GDT Items in the Whole Sample



Note. N = 1,028. GDT = Gambling Disorder Test; Corr = Correlation.
*** $p < .001$.

in assessing the GD construct. Notably, the highest association observed was between the GDT Item 1 and the GDT Item 3 ($r = .841$, $p < .001$) and the lowest association was between the GDT Item 3 and the GDT Item 4 ($r = .716$, $p < .001$).

The cross-correlations between the items of the GDT and PGSI are shown in Figure 2. As can be seen, these ranged from moderate to high in size and were all statistically significant at $p < .001$. Moreover, the lowest cross-correlation observed was between GDT Item 3 and PGSI Item 4 ($r = .500$, $p < .001$) and the highest between GDT Item 1 and PGSI Item 5 ($r = .778$, $p < .001$).

These robust inter-item and cross-item correlations suggest that the GDT consistently captures the essential components of GD, reflecting a cohesive and reliable measure that aligns well with the underlying diagnostic criteria, indicating that the GDT is a promising instrument for the assessment of GD symptoms.

Factorial Validity and Population Cross-Validity

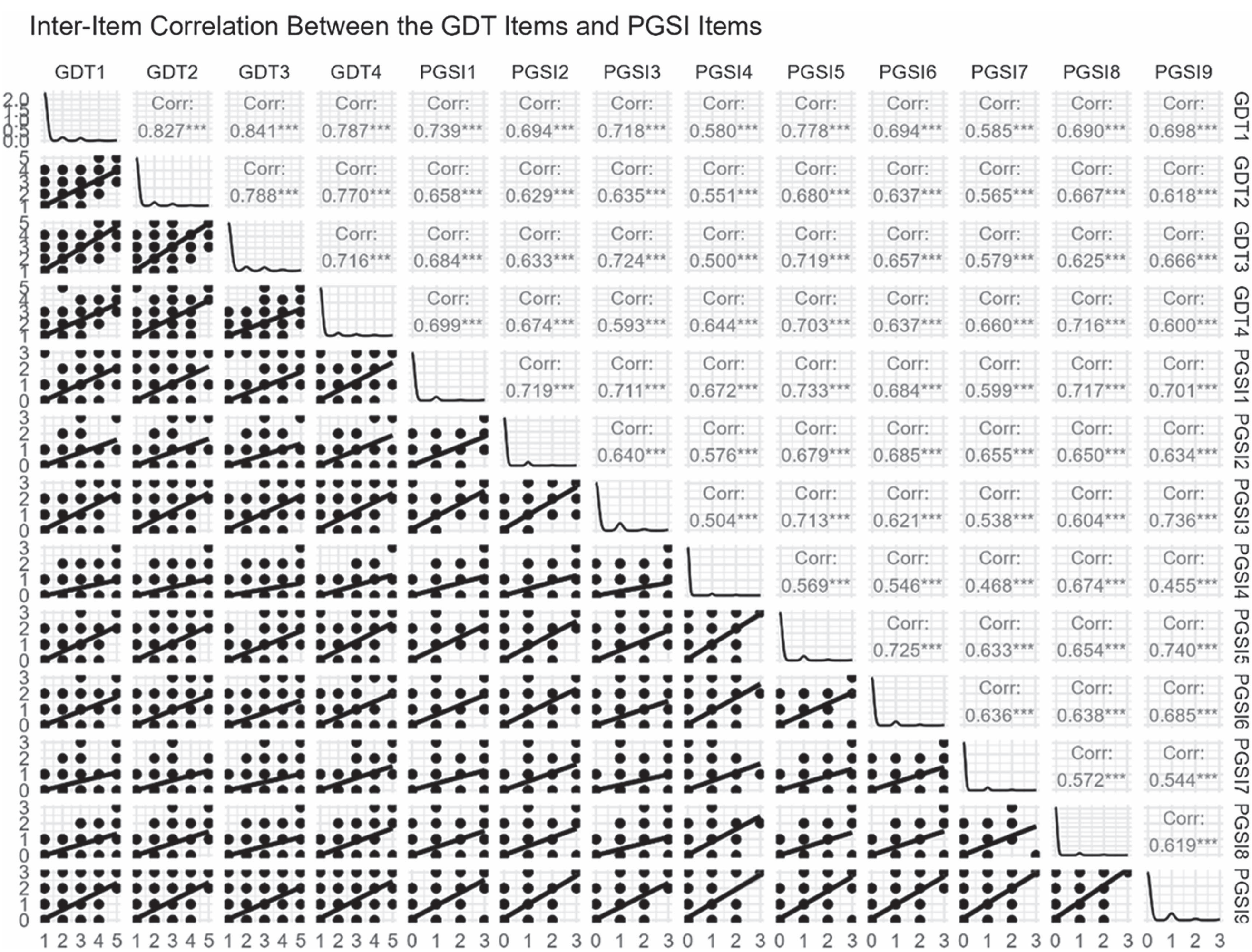
The next step of the psychometric analyses involved conducting an EFA to explore the factor structure of the GDT. The suitability of the data for EFA was examined with the Kaiser–Meyer–Olkin measure and Bartlett's Test of Sphericity. The Kaiser–Meyer–Olkin measure was .84, which supports the adequacy for factor analysis and the Bartlett's Test of Sphericity was significant, $\chi^2(6) = 1698.5$, $p < .001$, further supporting the suitability of the data for EFA.

Following this, the initial underlying factor structure of the GDT was examined via EFA in Sample 1 ($n = 514$) using Principal Axis Factoring with Direct Oblimin rotation (see Table 2). The results indicated that only the first factor reached an eigenvalue greater than one. Therefore, a one-factor solution was adopted as recommended by the Kaiser criterion (Kaiser, 1960). Overall, the unidimensional model solution showed acceptable fit, $\chi^2(2) = 12.13$, $p < .0023$; CFI = .994, TLI = .982, RMSEA = .099, 90% CI [.051, .156], and explained 76.52% of the total variance in the data. Furthermore, all standardized factor loadings (λ_{ij}) were considerably high (i.e., $\lambda_{ij} \geq .50$; Ferguson & Cox, 1993), suggesting a strong relationship between the observed items and the underlying latent factor extracted.

Additionally, as can be seen in Figure 3, the one-factor solution was further supported by the scree plot (Cattell, 1966) in combination with Horn's Parallel Analysis (Horn, 1965) which compared

Figure 2

Inter-Item Correlation Analysis Between All GDT and PGSI Items in the Whole Sample



Note. $N = 1{,}028$. GDT = Gambling Disorder Test; Corr = Correlation; PGSI = Problem Gambling Severity Index.
*** $p < .001$.

Table 2
Summary of the Results From the Exploratory Factor Analysis in Sample 1 (n = 514)

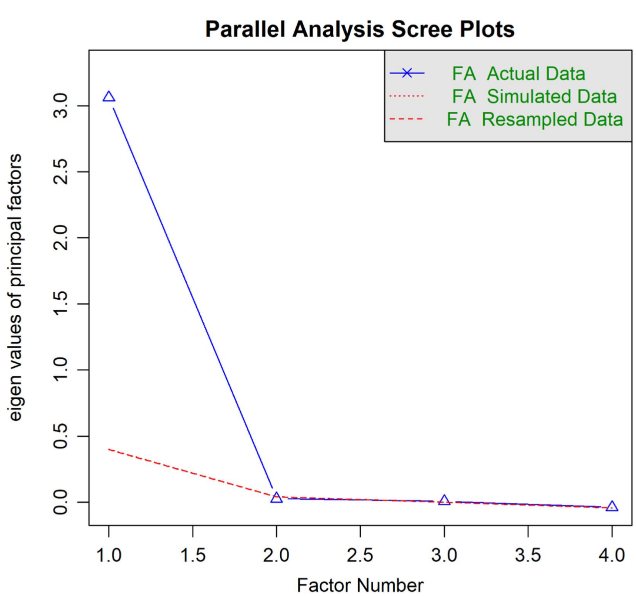
Item	λ_{ij}	Final communality	Eigenvalue Extracted	Eigenvalue Expected
1	.944	.890	3.061	3.288
2	.868	.753	0.028	0.329
3	.886	.785	0.008	0.251
4	.795	.632	−0.036	0.133

Note. Item wording was omitted from the table for the sake of clarity. For a full description of the items, please see Supplemental Table S1. A total of five iterations took place. Eigenvalue for the extracted factor = 3.29. The reported factor loadings are standardized. The sum of the communalities after the final extraction was 3.061. The expected eigenvalues were obtained from the parallel analysis. λ = Standardized factor loadings.

the eigenvalues of the actual data to those from 10,000 randomly generated simulated and resampled data sets.

Following this step, a CFA in Sample 2 ($n = 514$) was performed using the robust maximum likelihood estimator to further test the one-factor solution found in the previous step. The results of the analysis showed an adequate fit to the data, $\chi^2(2) = 10.031$, $p = .007$, CFI = .975, TLI = .925, RMSEA = .088, 90% CI [.067, .111], with $p = .002$ and SRMR = .020. With regard to the standardized factor loadings, these were all above the conventional threshold of $\lambda_{ij} \geq .50$ and statistically significant (see Table 3). To further test these findings, a second CFA model was estimated using the data from the complete sample ($N = 1{,}028$), with the results suggesting similar findings and further supporting the factorial validity of the GDT, $\chi^2(2) = 4.074$, $p = .130$, CFI = .996, TLI = .989, RMSEA = .032, 90% CI [.014, .049], with $p = .963$ and SRMR = .012. Taken together, the findings from the EFA and CFA support the factorial validity and population cross-validity of the GDT because a clear and interpretable one-factor solution with all items loading highly onto the latent factor was found across all models tested.

Figure 3
Screen Plot and Horn's Parallel Analysis Results Based on 10,000 Randomly Generated Simulated and Resampled Data Sets



Note. FA = Principal Axis Factor Analysis. See the online article for the color version of this figure.

Internal Consistency

The internal consistency of the GDT was tested using different reliability coefficients to ensure its robustness. As can be seen in Table 4, all coefficients (i.e., α, Ω, CR, and FD) that were calculated for the complete sample, Sample 1, and Sample 2 were high and above the expected thresholds. In sum, the findings suggest that the GDT showed excellent internal consistency levels.

Gender-Based Measurement Invariance

Due to the low number of individuals that reported their gender as "Other" (0.88%, $n = 9$), the measurement invariance analysis of the GDT was conducted for only female (50.97%, $n = 524$) and male (48.15%, $n = 495$) participants. The results of the multigroup CFA showed that the configural model provided a satisfactory fit, indicating that the factor structure of the GDT was consistent across females and males. Moreover, the metric invariance model tested suggested that the factor loadings were equivalent across genders, with the changes in both CFI (ΔCFI = −.006) and RMSEA (ΔRMSEA = −.001) being within acceptable thresholds, therefore supporting metric invariance. As for the scalar invariance model, the results indicated that the intercepts were equivalent across both genders, with changes in both CFI (ΔCFI = .000) and RMSEA (ΔRMSEA = −.001) situating within acceptable thresholds (see Table 5).

Because full scalar measurement was achieved, latent mean comparison was conducted to test whether female and male individuals differed in terms of GD symptoms, with males being the reference group (i.e., latent mean factor fixed to zero for identification purposes). The results showed that females scored significantly lower on the latent factor compared to males (latent mean difference = −0.156; $z = -3.844$, $p < .001$, $d = -.249$). This finding implies that males show a significantly higher latent mean compared to females, suggesting that they experience higher levels of GD symptoms than females as assessed by the GDT.

Further Validity Testing

As can be seen in Table 6, GDT total scores were positively associated with relevant measures of interest in the expected direction. More specifically, GDT total scores were positively associated with PGSI total scores ($r = .865$, $p < .0001$), experiencing of a significant problem in the past 12 months due to gambling ($r = .609$, $p < .0001$) and receiving complaints from others due to their gambling behaviors ($r = .533$, $p < .0001$). Overall, these results support both the criterion and concurrent validity of the GDT and further complement the inter-item and cross-item correlational analyses findings.

Table 3
Summary of the Results From the Confirmatory Factor Analysis in Sample 1 (n = 514) and Complete Sample (N = 1,028)

Item	λ_{ij}	*SE*	*z*	*p*	σ^2_ε	*SE*	*z*	*p*
				Sample 1 (n = 514)				
1	.932	N/A	N/A	N/A	.060	.013	4.715	<.001
2	.913	.059	15.093	<.001	.063	.019	3.238	<.001
3	.879	.065	15.625	<.001	.119	.022	5.396	<.001
4	.882	.053	15.195	<.001	.075	.016	4.777	<.001
Model fit	$\chi^2(2)$ = 10.031, p = .007, CFI = .975, TLI = .925, RMSEA = .088, 90% CI [.067, .111], p = .002; SRMR = .020							
				Complete sample 1 (N = 1,028)				
1	.939	N/A	N/A	N/A	.052	.009	5.757	<.001
2	.889	.044	18.779	<.001	.071	.013	5.537	<.001
3	.886	.044	23.393	<.001	.111	.016	7.081	<.001
4	.838	.045	16.605	<.001	.091	.015	5.872	<.001
Model fit	$\chi^2(2)$ = 4.074, p = .130, CFI = .996, TLI = .989, RMSEA = .032, 90% CI [.014, .049], p = .963; SRMR = .012							

Note. The σ^2_ε for the latent factor in Sample 1 was .397 (SE = .071, z = 5.6253, p < .001) and in the complete sample was .388 (SE = .051, z = 7.643, p < .001). N/A = not available due to item having its factor loading fixed; *SE* = standard error; λ = standardized factor loadings; σ^2_ε = nonstandardized item-level error variance; CFI = comparative fit index; TLI = Tucker–Lewis fit index; RMSEA = root-mean-square error of approximation; CI = confidence interval; SRMR = standardized root-mean-square residual.

Discussion

The present study sought to contribute to the field by developing the GDT, a brief standardized psychometric instrument for GD based on the latest WHO *ICD-11* diagnostic framework to ensure that the most recent clinical and diagnostic developments are adequately taken into account when assessing GD. To achieve this, a large nationally representative sample was recruited to aid the development of the GDT through several rigorous steps that aimed to scrutinize its psychometric properties applying both a CTT and IRT approach (see Supplemental Material).

Within the sample recruited, both participation in gambling activities and the prevalence rates of GD were similar to other studies. Within the U.K. adult population, the most recent Gambling Survey for Great Britain reported that approximately 48% of the population participated in gambling activities in the past 4 weeks (Gambling Commission, 2024), which is similar to the participation rate found in the present sample (i.e., 48.44%). In relation to the prevalence rates of GD in the present study, the self-reported prevalence found was 0.49%. This figure is relatively low and aligned with findings from other large and representative studies conducted in the United Kingdom where national statistics reported prevalence rates of GD of 0.40% in England (Gambling Commission, 2023) and Scotland (Gambling Commission, 2022) and 0.70% in Wales (Gambling Commission, 2018). In other culturally close countries to the United Kingdom, national estimates among adult populations have been reported to be around 0.30% in the Republic of Ireland (Mongan et al., 2022), 0.60% in Canada (Williams et al., 2021), and 0.60% in Australia (Gainsbury et al., 2014). The robust and close alignment between the prevalence rates of GD estimated with the GDT and other studies suggests that the newly developed instrument is capable of replicating epidemiological findings. However, it is important to note that the prevalence rate in the present study was not based on a clinical assessment but on self-reported responses to the GDT items. Therefore, this figure may differ from rates derived from clinical interviews or diagnostic evaluations conducted in a clinical setting.

Table 4
Summary of the Reliability Findings for Sample 1 (n = 514), Sample 2 (n = 514), and Complete Sample (N = 1,028)

Sample	α	Ω	Composite reliability	Factor determinacy
1	.924	.940	.928	.874
2	.942	.965	.946	.902
Complete	.934	.948	.938	.889

Note. α = Cronbach's alpha; Ω = McDonald's Omega.

Inter-item correlation analysis among the four new items of the GDT, as well as cross-item correlation analysis between the GDT and PGSI items were conducted to further support the psychometric evaluation of the GDT. These analyses explored how well the GDT items assessed GD and helped determining their satisfactory alignment with the PGSI. Because the GDT is grounded on the current WHO diagnostic framework for GD, it will aid expanding the current conceptual and diagnostic landscape, adding unique value to future clinical and research efforts to the assessment of GD.

In terms of validity of the GDT, the psychometric analyses yielded results supporting a unidimensional factor solution, with statistically significant and strong factor loadings. More specifically, the results of the EFA and CFA further supported the factorial validity and population cross-validity of the GDT scores. As for the reliability of the GDT, the results supported the scale's test score internal consistency through several coefficients across the complete sample and the two independent subsamples. The results of the further validity testing indicated that both criterion and concurrent validity were supported through the observed statistically significant positive associations between GDT total scores and (a) PGSI total scores, (b) the experience of a significant problem due to gambling, and (c) the receiving of complaints from others due to gambling behaviors. These findings mirror the literature showing that higher levels of GD are typically associated greater levels of PG symptoms (Stinchfield et al., 2016), the experience of functional impairments related to gambling (Sleczka & Romild, 2021), and problems

Table 5
Fit Indices for the Measurement Invariance Testing of the GDT Between Female and Male Genders

Invariance model	χ^2	*df*	*p*	CFI	TLI	RMSEA [90% CI]	SRMR	ΔCFI	ΔRMSEA	Decision
Configural	4.326	4	.364	.999	.998	.013 [.000, .035]	.012			Accept
Metric	7.724	7	.358	.998	.997	.014 [.000, .032]	.044	−.001	.001	Accept
Scalar	10.816	10	.372	.998	.998	.013 [.000, .030]	.044	.000	−.001	Accept

Note. Female (Group 1), $n = 524$. Male (Group 2), $n = 495$. GDT = Gambling Disorder Test; *df* = degrees of freedom; CFI = comparative fit index; TLI = Tucker–Lewis fit index; RMSEA = root-mean-square error of approximation; CI = confidence interval; SRMR = standardized root-mean-square residual.

between the gambler with concerned significant others such as family members (Tulloch, Browne, et al., 2021) and friends (Tulloch, Hing, et al., 2021) who may complain to the gambler about their behavior.

Another goal of the present study was to assess the measurement invariance of the GDT between male and female participants to better understand the scale's psychometric functioning between genders. Overall, the results supported the scale's configural, metric, and scalar invariance. The implications of these findings signify that at the configural level, both genders conceptualized the latent construct assessed (i.e., GD) in the same way as the unidimensional factor solution was identified for both male and female participants, suggesting that the basic structure of the scale is understood similarly across genders. At the metric level, the factor loadings were equivalent across both genders, which means that the items of the GDT contributed to the underlying latent construct to the same extent for both males and females. Finally, at the scalar level, the item intercepts were found to be equivalent across both genders, suggesting that differences in the test scores can be attributed to differences in the underlying latent construct rather than to differences in the way the two groups understood or responded to the items of the GDT.

Because full scalar invariance was achieved, latent mean comparison was conducted to provide further insight on gambling-related gender differences in the sample. Interestingly, the results of the latent mean comparison found that male participants scored significantly higher on the latent construct (i.e., GD) than female participants. This finding aligns with a previous systematic review study reporting that males are typically more likely than females to experience GD (Merkouris et al., 2016). Further research carried out in the general population (e.g., Husky et al., 2015) and clinical settings (Granero et al., 2009) also reported higher levels of GD among males while other studies have reported that males typically engage in gambling activities much earlier than females (Ronzitti et al., 2016).

Notwithstanding this, gender differences in GD are typically more nuanced and not as straightforward. For example, some studies have reported that females show higher GD prevalence than males (Håkansson & Widinghoff, 2020). These differences might be due to individual differences underpinning gambling behavior, activity, and underlying motives. For instance, one recent study reported that gambling-related gender profiles varied across genders, with females gambling for money significantly more than males (Hagfors et al., 2022), further illustrating the impact of gender differences in gambling activities.

Implications and Limitations

The findings obtained and the development of the GDT have important implications. At the clinical level, the development of such a brief and convenient screening instrument will likely have the potential to improve future clinical practices in the assessment of GD because many health services around the world have limited time and resources and may only be able to administer very brief screening instruments when assessing potential cases of GD (Dowling et al., 2019). At the research level, the GDT may facilitate further larger scale studies on GD because brief screening instruments showing high validity and reliability are increasingly needed to provide a convenient and cost-effective assessment of GD, particularly in population-level epidemiological studies (Dowling et al., 2019). It is worth noting that the findings in the present study support the use of the GDT across both male and female populations because the scale is able to function adequately when assessing GD among male and female gamblers.

Table 6
Overall Associations Between the GDT Total Scores and Other Relevant Variables

Variable	1	2	3	4	5	6	7
1. Gambling disorder (GDT)	—						
2. Problem gambling (PGSI)	.865	—					
3. Average monthly amount spent gambling	.360	.393	—				
4. Average hours per week spent gambling	.212	.203	.209	—			
5. Perceived level of engagement with gambling	.498	.443	.321	.300	—		
6. Complaints from others due to gambling behaviors	.533	.606	.267	.098	.336	—	
7. Significant problem experienced due to gambling	.609	.654	.340	.116	.364	.455	—

Note. All correlation coefficients statistically significant at $p < .05$. GDT = Gambling Disorder Test; PGSI = Problem Gambling Severity Index.

Although the study reported insightful findings, it is important to acknowledge its potential limitations. First, the GDT was developed based on data collected from the general population of a single country. Furthermore, the clinical utility of the scale warrants further testing in treatment-seeking populations to ensure that the robust psychometric properties found here can be also replicated in clinical studies. On this note, future clinical research will be needed to test the psychometric properties of the GDT further and to aid the development of cutoff points to assist in better distinguishing between disordered and nondisordered individuals. Second, another important limitation was that additional forms of validity, such as predictive validity, were not assessed. Future research will be needed to further examine whether the GDT can effectively predict key outcomes commonly associated with GD in order to strengthen its external validity and clinical applicability. For example, future studies might investigate how high GDT scores associate with financial difficulties, interpersonal conflicts, and/or legal issues emerging from disordered gambling behaviors, which would provide valuable insight into the instrument's utility in identifying at-risk individuals. Additionally, future research could explore the GDT's diagnostic concordance with clinical interview-based diagnoses to establish its predictive accuracy in real-world settings. Third, all findings from the present study were based exclusively on self-reported data, introducing the potential for biases, such as social desirability bias and short-term recall bias, which may influence participants' responses. These limitations are inherent to studies using psychometric instruments to assess latent constructs such as GD. However, employing more objective (e.g., behavioral) data or corroborative reports in future research could mitigate these potential biases and shed further light on the validity and reliability of the GDT scores, particularly among diverse populations and cultural contexts.

Finally, it is important to recognize that further research is needed to evaluate the GDT's score validity, particularly in relation to predictive utility relative to other types established GD criteria and standardized instruments. While the present study serves as an initial psychometric validation to establish the foundational psychometric properties of the GDT, future studies are needed to examine the instrument's unique contribution within a broader framework by assessing its ability to predict relevant external criteria because expanding the scope of the analyses to include additional comparative standardized tests beyond the PGSI would allow for a more comprehensive assessment of the GDT's structural positioning among existing GD measures and clarify its potential as a complementary or standalone diagnostic instrument.

Conclusion

The present study provides robust empirical support for the conceptualization and assessment of GD as defined in the WHO diagnostic framework, which refers to the diagnostic criteria for GD within the *ICD-11*. The findings obtained support the initial validity, reliability, and gender invariance of the GDT as a psychometric instrument for GD both within a CTT and IRT framework. Furthermore, the results demonstrated that the GDT is suitable for GD research requiring a convenient, concise, and updated assessment instrument for GD that reflects the latest advancements in the field.

The GDT represents an important contribution to the literature by offering a streamlined yet comprehensive psychometric approach reflecting the current understanding of GD. The GDT is a robust, updated assessment instrument capable of supporting clinicians and researchers alike in measuring relevant GD symptoms with adequate levels of reliability and validity, therefore expanding the range of existing screening instruments available for diverse clinical and research needs.

References

Abbott, M. W. (2020). Gambling and gambling-related harm: Recent World Health Organization initiatives. *Public Health*, *184*, 56–59. https://doi.org/10.1016/j.puhe.2020.04.001

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders* (5th ed., text rev.). https://doi.org/10.1176/appi.books.9780890425787

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*(2), 238–246. https://doi.org/10.1037/0033-2909.107.2.238

Bentler, P. M., & Bonnet, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, *88*(3), 588–606. https://doi.org/10.1037/0033-2909.88.3.588

Bernaards, C. A., & Jennrich, R. I. (2005). Gradient projection algorithms and software for arbitrary rotation criteria in factor analysis. *Educational and Psychological Measurement*, *65*(5), 676–696. https://doi.org/10.1177/0013164404272507

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

Caler, K., Garcia, J. R. V., & Nower, L. (2016). Assessing problem gambling: A review of classic and specialized measures. *Current Addiction Reports*, *3*(4), 437–444. https://doi.org/10.1007/s40429-016-0118-7

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, *1*(2), 245–276. https://doi.org/10.1207/s15327906mbr0102_10

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, *14*(3), 464–504. https://doi.org/10.1080/10705510701301834

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Dowling, N. A., Merkouris, S. S., Dias, S., Rodda, S. N., Manning, V., Youssef, G. J., Lubman, D. I., & Volberg, R. A. (2019). The diagnostic accuracy of brief screening instruments for problem gambling: A systematic review and meta-analysis. *Clinical Psychology Review*, *74*, Article 101784. https://doi.org/10.1016/j.cpr.2019.101784

Ferguson, E., & Cox, T. (1993). Exploratory factor analysis: A users' guide. *International Journal of Selection and Assessment*, *1*(2), 84–94. https://doi.org/10.1111/j.1468-2389.1993.tb00092.x

Ferris, J., & Wynne, H. (2001). *The Canadian Problem Gambling Index: Final report*. Canadian Centre on Substance Abuse. Retrieved February 24, 2025, from https://www.greo.ca/Modules/EvidenceCentre/files/Ferris%20et%20al(2001)The_Canadian_Problem_Gambling_Index.pdf

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39–50. https://doi.org/10.1177/002224378101800104

Gabellini, E., Lucchini, F., & Gattoni, M. E. (2023). Prevalence of problem gambling: A meta-analysis of recent empirical research (2016–2022). *Journal of Gambling Studies*, *39*(3), 1027–1057. https://doi.org/10.1007/s10899-022-10180-0

Gainsbury, S. M., Russell, A., Hing, N., Wood, R., Lubman, D. I., & Blaszczynski, A. (2014). The prevalence and determinants of problem gambling in Australia: Assessing the impact of interactive gambling and new technologies. *Psychology of Addictive Behaviors*, *28*(3), 769–779. https://doi.org/10.1037/a0036207

Gambling Commission. (2018). *Levels of problem gambling in Wales: Official statistics*. Retrieved January 18, 2025, from https://www.gamblingcommission.gov.uk/statistics-and-research/publication/levels-of-problem-gambling-in-wales

Gambling Commission. (2022). *Levels of problem gambling in Scotland: Official statistics*. Retrieved January 18, 2025, from https://www.gamblingcommission.gov.uk/statistics-and-research/publication/levels-of-problem-gambling-in-scotland

Gambling Commission. (2023). *Levels of problem gambling in England: Official statistics*. Retrieved January 18, 2025, from https://www.gamblingcommission.gov.uk/statistics-and-research/publication/levels-of-problem-gambling-in-england

Gambling Commission. (2024). *Statistics on gambling participation—Year 1 (2023), Wave 2: Official statistics*. Retrieved January 18, 2025, from https://www.gamblingcommission.gov.uk/statistics-and-research/publication/statistics-on-gambling-participation-year-1-2023-wave-2-official-statistics

Granero, R., Penelo, E., Martínez-Giménez, R., Álvarez-Moya, E., Gómez-Peña, M., Aymamí, M. N., Bueno, B., Fernández-Aranda, F., & Jiménez-Murcia, S. (2009). Sex differences among treatment-seeking adult pathologic gamblers. *Comprehensive Psychiatry*, *50*(2), 173–180. https://doi.org/10.1016/j.comppsych.2008.07.005

Hagfors, H., Castrén, S., & Salonen, A. H. (2022). How gambling motives are associated with socio-demographics and gambling behavior—A Finnish population study. *Journal of Behavioral Addictions*, *11*(1), 63–74. https://doi.org/10.1556/2006.2022.00003

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis: A global perspective* (7th ed.). Pearson Prentice Hall.

Håkansson, A., & Widinghoff, C. (2020). Gender differences in problem gamblers in an online gambling setting. *Psychology Research and Behavior Management*, *13*, 681–691. https://doi.org/10.2147/PRBM.S248540

Harrell, F. E., Jr., & Dupont, C. (2024). *Hmisc: Harrell miscellaneous* [Computer software]. https://doi.org/10.32614/CRAN.package.Hmisc

Ho, D., Imai, K., King, G., & Stuart, E. A. (2011). MatchIt: Nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software*, *42*(8), 1–28. https://doi.org/10.18637/jss.v042.i08

Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural equation modelling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods*, *6*(1), 53–60. https://academic-publishing.org/index.php/ejbrm/article/view/1224

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*(2), 179–185. https://doi.org/10.1007/BF02289447

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Husky, M. M., Michel, G., Richard, J. B., Guignard, R., & Beck, F. (2015). Gender differences in the associations of gambling activities and suicidal behaviors with problem gambling in a nationally representative French sample. *Addictive Behaviors*, *45*, 45–50. https://doi.org/10.1016/j.addbeh.2015.01.011

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, *20*(1), 141–151. https://doi.org/10.1177/001316446002000116

Kim, H. S., & Hodgins, D. C. (2019). A review of the evidence for considering Gambling Disorder (and other behavioral addictions) as a disorder due to addictive behaviors in the *ICD-11*: A focus on case–control studies. *Current Addiction Reports*, *6*(3), 273–295. https://doi.org/10.1007/s40429-019-00256-0

Langham, E., Thorne, H., Browne, M., Donaldson, P., Rose, J., & Rockloff, M. (2015). Understanding gambling related harm: A proposed definition, conceptual framework, and taxonomy of harms. *BMC Public Health*, *16*(1), Article 80. https://doi.org/10.1186/s12889-016-2747-0

Lesieur, H. R., & Blume, S. B. (1987). The South Oaks Gambling Screen (SOGS): A new instrument for the identification of pathological gamblers. *The American Journal of Psychiatry*, *144*(9), 1184–1188. https://doi.org/10.1176/ajp.144.9.1184

Merkouris, S. S., Thomas, A. C., Shandley, K. A., Rodda, S. N., Oldenhof, E., & Dowling, N. A. (2016). An update on gender differences in the characteristics associated with problem gambling: A systematic review. *Current Addiction Reports*, *3*(3), 254–267. https://doi.org/10.1007/s40429-016-0106-y

Mongan, D., Millar, S. R., Doyle, A., Chakraborty, S., & Galvin, B. (2022). *Gambling in the Republic of Ireland: Results from the 2019–20 National Drug and Alcohol Survey*. Health Research Board. Retrieved February 24, 2025, from https://www.drugsandalcohol.ie/35305/

Müller, S. M., Wegmann, E., Oelker, A., Stark, R., Müller, A., Montag, C., Wölfling, K., Rumpf, H. J., & Brand, M. (2022). Assessment of Criteria for Specific Internet-use Disorders (ACSID-11): Introduction of a new screening instrument capturing *ICD-11* criteria for gaming disorder and other potential internet-use disorders. *Journal of Behavioral Addictions*, *11*(2), 427–450. https://doi.org/10.1556/2006.2022.00013

Muthén, L. K., & Muthén, B. O. (2018). *Mplus user's guide* (8th ed.).

Otto, J. L., Smolenski, D. J., Garvey Wilson, A. L., Evatt, D. P., Campbell, M. S., Beech, E. H., Workman, D. E., Morgan, R. L., O'Gallagher, K., & Belsher, B. E. (2020). A systematic review evaluating screening instruments for gambling disorder finds lack of adequate evidence. *Journal of Clinical Epidemiology*, *120*, 86–93. https://doi.org/10.1016/j.jclinepi.2019.12.022

Pontes, H. M., & Griffiths, M. D. (2015). Measuring *DSM-5* internet gaming disorder: Development and validation of a short psychometric scale. *Computers in Human Behavior*, *45*, 137–143. https://doi.org/10.1016/j.chb.2014.12.006

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, *41*, 71–90. https://doi.org/10.1016/j.dr.2016.06.004

R Core Team. (2024). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org

Raju, N. S., Bilgic, R., Edwards, J. E., & Fleer, P. F. (1997). Methodology review: Estimation of population validity and cross-validity, and the use of equal weights in prediction. *Applied Psychological Measurement*, *21*(4), 291–305. https://doi.org/10.1177/01466216970214001

Raju, N. S., Bilgic, R., Edwards, J. E., & Fleer, P. F. (1999). Accuracy of population validity and cross-validity estimation: An empirical comparison of formula-based, traditional empirical, and equal weights procedures. *Applied Psychological Measurement*, *23*(2), 99–115. https://doi.org/10.1177/01466219922031220

Reilly, C., & Smith, N. (2013). *The evolving definition of pathological gambling in the DSM-5*. National Center for Responsible Gaming.

Revelle, W. (2024). *psych: Procedures for psychological, psychometric, and personality research* (R package Version 2.4.3) [Computer software]. Northwestern University. https://CRAN.R-project.org/package=psych

Ronzitti, S., Lutri, V., Smith, N., Clerici, M., & Bowden-Jones, H. (2016). Gender differences in treatment-seeking British pathological gamblers. *Journal of Behavioral Addictions*, *5*(2), 231–238. https://doi.org/10.1556/2006.5.2016.032

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Saunders, J. B., Degenhardt, L., & Farrell, M. (2017). Excessive gambling and gaming: Addictive disorders? *The Lancet Psychiatry*, *4*(6), 433–435. https://doi.org/10.1016/S2215-0366(17)30210-9

Schloerke, B., Cook, D., Larmarange, J., Briatte, F., Marbach, M., Thoen, E., Elberg, A., & Crowley, J. (2024). *GGally: Extension to "ggplot2"* (R package Version 2.2.1) [Computer software]. https://github.com/ggobi/ggally

Sirola, A., Nyrhinen, J., & Wilska, T. A. (2023). Psychosocial perspective on problem gambling: The role of social relationships, resilience, and COVID-19 worry. *Journal of Gambling Studies*, *39*(3), 1467–1485. https://doi.org/10.1007/s10899-022-10185-9

Sleczka, P., & Romild, U. (2021). On the stability and the progression of gambling problems: Longitudinal relations between different problems related to gambling. *Addiction*, *116*(1), 116–125. https://doi.org/10.1111/add.15093

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

Stinchfield, R., McCready, J., Turner, N. E., Jimenez-Murcia, S., Petry, N. M., Grant, J., Welte, J., Chapman, H., & Winters, K. C. (2016). Reliability, validity, and classification accuracy of the *DSM-5* diagnostic criteria for gambling disorder and comparison to *DSM-IV*. *Journal of Gambling Studies*, *32*(3), 905–922. https://doi.org/10.1007/s10899-015-9573-7

Tulloch, C., Browne, M., Hing, N., & Rockloff, M. (2021). The relationship between family gambling problems, other family stressors, and health indicators in a large population-representative sample of Australian adults. *Journal of Gambling Studies*, *37*(4), 1139–1162. https://doi.org/10.1007/s10899-020-09990-x

Tulloch, C., Hing, N., Browne, M., Rockloff, M., & Hilbrecht, M. (2021). The effect of gambling problems on the subjective wellbeing of gamblers' family and friends: Evidence from large-scale population research in Australia and Canada. *Journal of Behavioral Addictions*, *10*(4), 941–952. https://doi.org/10.1556/2006.2021.00077

Wardle, H., Reith, G., Langham, E., & Rogers, R. D. (2019). Gambling and public health: We need policy action to prevent harm. *The BMJ*, *365*, Article l1807. https://doi.org/10.1136/bmj.l1807

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer. https://doi.org/10.1007/978-3-319-24277-4

Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2023). *dplyr: A grammar of data manipulation* (R package Version 1.1.4) [Computer software]. https://github.com/tidyverse/dplyr

Wickham, H., Vaughan, D., Girlich, M., & Ushey, K. (2024). *tidyr: Tidy Messy data* [Computer software]. https://doi.org/10.32614/CRAN.package.tidyr

Williams, R. J., Leonard, C. A., Belanger, Y. D., Christensen, D. R., El-Guebaly, N., Hodgins, D. C., McGrath, D. S., Nicoll, F., & Stevens, R. M. G. (2021). Gambling and problem gambling in Canada in 2018: Prevalence and changes since 2002. *The Canadian Journal of Psychiatry*, *66*(5), 485–494. https://doi.org/10.1177/0706743720980080

World Health Organization. (2022a). *6C50 gambling disorder*. Retrieved June 20, 2024, from https://icd.who.int/browse/2024-01/mms/en#1041487064

World Health Organization. (2022b). *International statistical classification of diseases for mortality and morbidity statistics* (11th rev.). Retrieved January 18, 2025, from https://icd.who.int/

Received July 8, 2024
Revision received January 24, 2025
Accepted January 29, 2025 ■