Journal of Gambling Issues

Open Access Original Research

The 20-item Gambling Harms Scale (GHS-20): Benchmarked to health utility using propensity weighting and control for comorbidities

Matthew Browne[1,3], Vijay Rawat[1], Matthew J Rockloff[1,4], Alex MT Russell[1], Nerilee Hing[1,5], En Li[1], Georgia Dellosa[1], Philip Newall[1,2,6],

Citation: Browne, M., Rawat, V., Rockloff, M.J., Russell1, A.M.T., Hing, N., Li, E., Dellosa, G., Newall1, P. (2025). The 20-item Gambling Harms Scale (GHS-20): Benchmarked to health utility using propensity weighting and control for comorbidities. Journal of Gambling Issues.

Editor-in-Chief: Nigel Turner, Ph.D.

Received: 05/25/2025
Accepted: 10/29/2025
Published: 11/05/2025

[1]Experimental Gambling Research Laboratory, Central Queensland University, Australia
[2]School of Psychological Science, Bristol University, U.K.
[3]**ORCiD**: 0000-0002-2668-6229
[4]**ORCiD**: 0000-0002-0080-2690
[5]**ORCiD**: 0000-0002-0466-5611
[6]**ORCiD**: 0000-0002-1660-9254
*Corresponding author: Matthew Browne: m.browne@cqu.edu.au

Abstract. Measuring gambling-related harm is crucial for public health initiatives, but existing screens like the Problem Gambling Severity Index have limitations. The 10-item Short Gambling Harms Screen addresses some gaps, but lacks comprehensive domain coverage and robust validation against health utility benchmarks. This study aimed to develop and validate an extended 20-item Gambling Harms Scale. Objectives included achieving better representation across harm domains, selecting items based on unique associations with health utility decrements, and benchmarking the scale's scores against health utility. Data were collected from 2,603 Australian adults who gambled in the past year. Participants completed 31 candidate harm items, measures of gambling problems, quality of life, and health functioning. Item selection utilised latent trait modelling and lasso regression. The final scale was benchmarked against health outcomes using propensity score weighting to balance demographic factors and generalised additive modelling to control for comorbidities. Lasso regression identified items providing unique explanatory information, resulting in the 20-item harm scale, which includes items from all six gambling harm domains. The extended scale demonstrated excellent reliability (alpha=.98, omega=.90), correlated positively with the PGSI (r=.78) and negatively with health utility (r=-.38), but statistically was not a meaningful improvement on the 10-item scale. GAM analysis revealed a significant, near-linear negative relationship between harm scores and health utility, with statistically significant decrements observed even at low scores. Like the shorter scale, the extended scale is psychometrically robust, and can comprehensively assess gambling harm and calculate population harm burden.

Keywords: Public Health, Gambling-related Harm, Population Harm, Health Impacts.

Introduction

In jurisdictions around the world with legal gambling, authorities have a mandated responsibility to reduce harm from gambling. This is in line with a public health approach to managing gambling, which treats the issue in an epidemiological framework like that applied to tobacco or alcohol (Wardle et al., 2024). The benchmark of health and wellbeing is central to this framework, and this is reflected in definitions of gambling harm that emphasise decrements to the same (Wardle et al., 2018, Langham et al., 2015). Although population screens such as the Problem Gambling Severity Index (PGSI; Ferris & Wynne, 2001) have advanced our understanding of gambling harms, they are not ideal measures for monitoring population levels of gambling harm. This is because they (a) conflate the harms from problem gambling (i.e. the outcome) with the risk factor of behavioural addiction, (b) yield only prevalence estimates (e.g. of "problem gamblers") rather than a total metric of impact, (c) are ambiguous about the status of "at risk" gamblers (i.e. whether they are harmed or not) and (d) provide no guidance as to how persons of varying risk or impact might be combined into a single population metric. In short, screens for problem gambling have only an indirect theoretical link to health and wellbeing (Browne, 2019), and were not designed with this benchmark in mind.

The Short Gambling Harms Screen (SGHS) (Browne et al., 2017) is a 10-item measure intended to address a gap in the measurement of gambling harm, which has a direct theoretical link to health impact (which we will henceforth refer to the GHS-10 here for reasons explained subsequently). Although several other approaches to gambling harm measurement exist (e.g., Gooding et al., 2024; Mou et al., 2025), the GHS-10 is the most widely adopted instrument, has been translated into several languages for use in multiple jurisdictions, and has received extensive psychometric validation (Browne et al., 2021). The Gambling Harm Measure (GHM; Gooding et al., 2024) has also demonstrated strong psychometric properties in large-scale survey validation; however, it is longer and less widely used to date. Nevertheless, concerns about the GHS-10 have been expressed in the literature (Delfabbro & King, 2017), and also (more commonly) at presentations given at gambling governance conference venues. Some concerns, such as the binary response format, or speculation that some items might reflect rational opportunity costs of gambling, have already been addressed empirically (Murray-Boyle et al., 2021; McLauchlan et al., 2020). However, there is still some uncertainty and debate as to whether the GHS-10 truly indexes meaningful impacts to health and wellbeing. Furthermore, due to its brevity, the GHS-10 does not provide complete coverage of all domains of gambling related harm: physical health, relationships, finances, psychological and emotional, work/study and social deviance (Langham et al., 2015). Furthermore, the

item selection process (from an initial pool of 72 items) for the GHS-10 was based on sensitivity (i.e. minimising false negative zero scores), rather than with respect to an external validation metric.

Both the public health framework and standard definitions of gambling harm advise for the use of *health utility* as the key outcome of interest and the ultimate yardstick with which to evaluate scores on a screen intended to encompass negative impact from gambling. Health utility refers to a standardized metric of health-related quality of life, scaled from 0 (death) to 1 (perfect health), and widely used in health economics (Brazier et al., 2017). Previous work has described how a public health measurement framework would be applied in order to estimate health decrements associated with gambling (Browne, Rawat, et al., 2020). That work outlined a process for indirect elicitation of decrements to *experienced health utility* associated with gambling, which avoids potential response bias introduced by preference-based elicitation using direct evaluation of vignettes or descriptions. Preference-based elicitation methods derive health utility values from individuals' stated preferences about health states, often through structured instruments such as the EQ-5D or SF-6D (Brazier & Deverill, 1999). Indirect elicitation, on the other hand, refers to the approach of mapping self-reported health states onto utility values using established scoring algorithms, rather than asking individuals to directly value those states (Brazier & Deverill, 1999). This involves collecting a sample of harmed and unharmed individuals, which is then weighted based on known socio-demographic risk factors, so as to balance the groups based on their propensity for experiencing harms or problems. Multivariate regression is then used to estimate the unique relationship between the screen and health utility, whilst controlling for known comorbidities identified from the literature (e.g. diagnosis with another mental disorder, alcohol use) that also impact health. Given that it is unfeasible and unethical to conduct a controlled experiment (that would involve 'giving' the treatment group gambling problems), this framework has clear advantages in estimating the direct link between gambling harm symptomatology and health utility.

The framework described above has previously been applied to the GHS-10 and the PGSI (Browne et al., 2022), using the SF-6D based on the SF-12 health survey (Burdine et al., 2000, Busija et al., 2011). This benchmark is a self-report assessment of health functioning, from which raw scores can be converted to a set of reference health utilities. The preference-based elicitation of these utilities' mean scores are ratio comparable, and decrements attributable to health conditions can in-principle be summed over individuals to yield an assessment of population-level impact. Previous work calculated implied utility decrements for both gambling screens, finding that all non-zero GHS-10 scores were associated with significant decrements to health utility (Browne et al., 2022). Significant decrements were also detected for PGSI moderate risk and problem gamblers, but not for low risk gamblers.

Based on development of the SGHS/GHS-10, we anticipated that extending the GHS-10 to 20 items would add little in terms of statistical performance. However, the inclusion of a 20-item version was motivated by stakeholder and funding body input, which emphasised the importance of face validity and coverage across all theorised domains of gambling harm (Langham et al., 2015). Accordingly, the GHS-20 was designed to complement the GHS-10: the shorter version provides brevity and efficiency, while the longer version offers broader domain coverage and acceptability among end-users. The current study therefore evaluates both versions, establishing their psychometric properties and benchmarking them against health utility.

Australia presents a particularly salient gambling environment, with the nation recording the highest per capita gambling losses globally at $32 billion annually in 2022-23 (Queensland Government Statistician's Office, 2024). Recent data demonstrates an alarming upward trend in both participation and harm, with gambling participation among Australian adults increasing from 57% in 2019 to 65.1% in 2024, while those experiencing gambling-related harms rose from 11% to 15% over the same period (Australian Institute of Family Studies, 2024). This escalating pattern of gambling involvement and associated harms, mostly from Electronic Gaming Machines (EGMs), wagering, and sports betting, illustrates the critical need for robust measurement tools to accurately assess and monitor gambling-related impacts across the Australian population.

Aims

The present study aimed to develop and validate an extended 'sister' scale to the GHS-10 in the Australian context, with the properties of:

a) better representation of all harm domains, as described in Langham et al. (2015)
b) more emphasis on moderate-to-severe harms, as suggested by Delfabbro and King (2017)
c) additional items selected based on unique associations with decrements to health utility
d) calculated implied health utility decrements for the total sum score

At the outset, we determined that the total length of the new scale should be 20 items, complementing the original 10-item harm scale. For the *causal regression* (see caveat below) stage of our measurement framework, we employ a generalised additive model (GAM) approach for estimating the relationship between both scales and health utility. This appears to be the most elegant approach to handle potential non-linearity in the relationship. Henceforth we will refer to the new scale as the Gambling Harms Scale 20 (GHS-20) and the original SGHS as the Gambling Harms Scale 10 (GHS-10). This naming scheme is adopted to ensure consistency with corresponding versions for affected others that are in-development. Since there is a law of diminishing returns regarding psychometric performance for longer scales, we leave as an open question as to whether the GHS-20

performs measurably better in psychometric terms than the GHS-10. Accordingly, our research questions were as follows:

a) Can a 20-item scale be developed that provides coverage across all six domains of gambling harm while maintaining strong psychometric properties?
b) Which gambling harm items provide unique explanatory information about health utility decrements when controlling for item redundancy?
c) What is the relationship between GHS-20 scores and health utility (SF-6D) after controlling for demographic confounders and comorbidities?
d) Do low levels of gambling harm (scores 1-2) show statistically significant health utility decrements, supporting the scale's sensitivity for population-level assessment?

Method

Ethical approval for this study was received from the Central Queensland University Ethics Committee (approval #22341) and all methods were performed in accordance with the relevant guidelines and regulations. Participants provided informed consent before participating.

Measures

Table 1 provides an index to the gambling harm items evaluated in this study, and their associated domains. The 10 GHS-10 items are marked with *, whereas items marked + are referred to in the results section and include those additional items that ultimately form the GHS-20. The selection process for these items is detailed below.

Table 1. Abbreviations, full item content, and domain categories of candidate gambling harm items

Abbreviation	Item	Domain
(1) Red. Spending *+	Reduction of my available spending money	Financial
(2) Red. Rec. Exp. *+	Less spending on recreational expenses such as eating out, going to movies or other entertainment	Financial
(3) Red. Savings *+	Reduction of my savings	Financial
(4) Sold Items *+	Sold personal items	Financial
(5) Inc. CC Debt *	Increased credit card debt	Financial
(6) Late Bills	Late payments on bills (e.g. utilities, rates)	Financial
(7) Red. Ess. Exp.+	Less spending on essential expenses such as medications, healthcare and food	Financial
(8) Time+	Used my work or study time to gamble	Work/Study

(9) Red. Perf.+	Reduced performance at work or study (i.e. due to tiredness or distraction)	Work/Study
(10) Absent+	Was absent from work or study	Work/Study
(11) Red. Sleep Gamb.	Loss of sleep due to spending time gambling	Health
(12) Malnutr.	Didn't eat as much or often as I should	Health
(13) Dep.+	Increased experience of depression	Health
(14) Regret *	Had regrets that made me feel sorry about my gambling	Emo/Psych
(15) Fail. *+	Felt like a failure	Emo/Psych
(16) Shame *	Felt ashamed of my gambling	Emo/Psych
(17) Distress *+	Felt distressed about my gambling	Emo/Psych
(18) Vulner.+	Felt insecure or vulnerable	Emo/Psych
(19) Worthl.+	Felt worthless	Emo/Psych
(20) Red. Time *	Spent less time with people I care about	Relationships
(21) Red. Events	Spent less time attending social events (non-gambling-related)	Relationships
(22) Inc. Tension	Experienced greater tension in my relationships (suspicion, lying, resentment, etc.)	Relationships
(23) Isolation+	Social isolation (felt excluded or shut-off from others)	Relationships
(24) Inc. Conflict+	Experienced greater conflict in my relationships (arguing, fighting, ultimatums)	Relationships
(25) Threat End.	Threat of separation or ending a relationship/s	Relationships
(26) Red. Contrib. Cult.	Reduced my contribution to religious or cultural practices	Soc. Deviance
(27) Pay Mon.+	Promised to pay back money without genuinely intending to do so	Soc. Deviance
(28) Viol.	Had experiences with violence (including family/domestic violence)	Soc. Deviance
(29) Chil. Negle.	Didn't fully attend to the needs of children	Soc. Deviance
(30) Took Mon.	Took money or items from friends or family without asking first	Soc. Deviance
(31) Theft Gov.	Petty theft or dishonesty in respect to government, business, or other people (not family/friends)	Soc. Deviance

Note: This table presents all 31 candidate gambling harm items evaluated in the study, organized by harm domain (Langham et al., 2015). Asterisks (*) denote items included in the original 10-item Short Gambling Harm Screen (SGHS, now termed GHS-10); plus signs (+) denote items identified in the present analysis for inclusion in the 20-item Gambling Harms Scale (GHS-20). Items marked with both symbols appear in both scales. The final GHS-20 comprises all items marked with +, resulting in coverage across all six harm domains.

Participants reported on their demographic characteristics such as age, gender, and country of birth (see Table 2 for key characteristics).

Problem gambling status was assessed using the Problem Gambling Severity Index (Ferris & Wynne, 2001). The PGSI uses nine items (e.g., "Have you bet more than you could really afford to lose?") measured on a

four-point scale from 0 (never) to 3 (almost always). All nine items are summed with resulting classifications of 0 (non-problem gambler), 1 to 2 (low-risk gambler), 3 to 7 (moderate-risk gambler), or 8+ (problem gambler). Coefficient alpha was .95, indicating excellent internal consistency, and coefficient omega was .92, which provides a more robust estimate by also accounting for the strength of factor loadings.

Gambling harms were assessed using a checklist of 31 items covering six domains of harm (financial [7-items], relationship [6], emotional/psychological [6], health [3], work/study [3], and social deviance [6]). These items were selected based on formative research (Browne & Rockloff, 2018; Browne, Volberg, et al., 2020) that identified which harms from the original 72-item checklist were at least moderately prevalent among those experiencing some degree of gambling problems. By 'at least moderately prevalent,' we refer to harms that, in our prior datasets (population-representative and internet-panel), showed non-trivial endorsement among harmed respondents across multiple samples. We did not impose a single numerical cut-off; instead, we excluded items that were consistently rare across studies (e.g., very low single-digit prevalence) or idiosyncratic to a single sample, while prioritising coverage across all harm domains. Any residual sample-specific differences in prevalence were further addressed at the multivariate selection stage, where items were retained only if they provided unique explanatory information about health utility and WHOQOL-BREF outcomes (lasso-based selection). The referenced studies comprised both internet panel and population-representative screens. Thus, in the interest of managing survey length, it was possible to exclude from consideration harms with very low prevalence (e.g. attempted suicide or loss of significant assets), which would entail they provide little information in a new harms screen. Certain items, such as 'increased use of tobacco' present issues in a health context, terms of assessing a comorbidity and were excluded based on content non-central to the experience of gambling harm. Items comprising the SGHS (now GHS-10) (Browne et al., 2017) comprised a subset of these items. Item abbreviations and full item content are presented in Table 1.

Quality of life was assessed using the 26-item WHOQOL-BREF (World Health Organization, 1996). Each item is measured on a five-point scale. Apart from the SF-6D (described below), the WHOQOL-BREF is probably the most well known and most-used instrument to assess global health and wellbeing. The measure comprises two general items about overall quality of life, as well as four overall domains including physical health (7 items), psychological (6 items), social relationships (3 items), and environment (8 items). Three items require reverse scoring. Scores for the four domains are transformed so that each domain's score ranges from 4 to 20, with higher total scores indicating higher perceived quality of life in that respective domain. A composite quality of life score was also calculated using equally

weighted average scores from the four domains to provide a single benchmark. However, consistent with recommendations by the instrument authors, the main analyses, including item selection, were conducted using the WHOQOL-BREF subscales.

Health utility was measured using the SF-6D (Brazier et al., 2004), which is a measure derived from the SF-12 self-report measure (Burdine et al., 2000). It involves conversion of the raw SF-12 responses to health utility scores using data from preference-based protocols, in which respondents are asked to consider how much life one is willing to sacrifice or risk, to avoid the limitation in functioning described. It captures physical functioning, role limitations, social functioning, pain, mental health, and vitality; and yields health utility coefficients between 0.345 to 1.000 (Hunger et al., 2011). These scores are intended to function as a true metric, in the sense that they are bounded by zero (impairment so severe that life is not worth living) to one (full health), and ratios of scores are consistent. That is, a decrement of 0.2 is twice as bad as a decrement of 0.1 since it implies that a typical person would give up (or risk) twice as much time to avoid a fixed amount of time living with the former condition. The SF-6D has been used in fields related to gambling, such as to estimate health utility for alcohol use disorder (Barbosa et al., 2021), but except for one previous study (Browne et al., 2022), has not hitherto been used to evaluate the impact of gambling-related harm.

Scale development

The Gambling Harms Scale 20 (GHS-20) was designed as an extension of the ten-item SGHS/GHS-10, following stakeholder and funder input that emphasised the importance of face validity and coverage across all theorised domains of harm (Langham et al., 2015). Candidate items were drawn from a 31-item harm checklist that itself had been distilled from a 72-item formative item pool. Items that were consistently rare were excluded, with priority given to ensuring representation across six harm domains: financial, relationship, emotional/psychological, health, work/study, and social deviance. Psychometric evaluation informed the final reduction to 20 items. Tetrachoric correlations, confirmatory factor analysis, and two-parameter logistic trait models were used to assess internal consistency, dimensionality, and item properties. Although subfactor structures were examined for theoretical interest, the scale was conceptualised and validated primarily as a unidimensional measure of gambling harm. Multivariate lasso regression was then applied to predict decrements in health-related quality of life (SF-6D and WHOQOL-BREF subscales), with coefficients constrained to be zero or negative so that retained items each reflected unique, robust associations with reduced quality of life. The resulting 20 items thus combined conceptual domain coverage with empirical evidence of predictive validity for health outcomes.

Participants

Participants were recruited from a commercial panel provider during late 2020 and early 2021. The commercial provider maintains a pool of pre-registered respondents who have signed up to receive survey invitations. Panel members earn points for participating, which can subsequently be redeemed for vouchers, goods, or cash, incentivising engagement and panel retention. As compensation, participants received points which could be exchanged for prizes and/or cash via their panel. To be eligible, participants were required to be Australian residents, aged 18 years or above, and to have gambled in the past 12 months. Gambling included participating in at least one of the following activities: race betting, electronic gaming machines (pokies), casino table games, sports betting, informal private betting for money, Keno, bingo, esports betting, and fantasy sports betting. Residents of the state of Victoria were excluded due to multiple Covid-19 lockdowns which meant no access to venue-based gambling, including at time of sampling. Using soft quotas, we aimed to sample approximately equal groups with respect to age (18-29, 30-44, 45+) and gender, and with respect to *reference* (i.e., no harms nominated) versus *affected* group status (> 1 harms on the GHS-10). Of the 22,699 respondents who began the survey, 16,061 were screened out. Of those who screened out, 5,848 did not meet the residency or age criteria, 5922 withdrew (i.e., did not complete the survey), 441 provided poor quality data such as straight lining through the survey, and 3850 were excluded due to quotas being full. A total of 6,638 responses were retained, of which 2,603 were people who gamble, who formed part of the present analysis. The remaining responses included a sample of those affected / unaffected by another's gambling, which did not form part of the present study. A total of 1,546 (59.4%) scored zero on the GHS-10 and 1,057 (40.6%) scored 1+. Table 2 provides the demographic characteristics for these gamblers who scored zero (the reference group) versus 1+ (the affected group) on the GHS-10.

Table 2. Descriptive characteristics for the sample of gamblers by GHS-10 status

Variable	**Reference**	**Affected**
	No harms nominated n (%)	**1+ harms nominated n (%)**
Total	**1546 (100)**	**1057 (100)**
Gender		
Male	880 (56.9)	613 (58.0)
Female	665 (43.0)	443 (41.9)
Other	1 (0.1)	1 (0.1)
Mean age (SD [years])	51.16 (17.48)	42.15 (16.03)
Country of birth		
Australia	1232 (79.7)	845 (79.9)

Other	314 (20.3)	212 (20.1)
Main language spoken at home		
English	1497 (96.8)	991 (93.8)
Other	49 (3.2)	66 (6.2)
Aboriginal or Torres Strait Islander origin		
No	1463 (94.6)	958 (90.6)
Yes	83 (5.4)	99 (9.4)
State/Territory of residence		
New South Wales	652 (42.2)	526 (49.8)
Queensland	452 (29.2)	258 (24.4)
South Australia	196 (12.7)	119 (11.3)
Tasmania	56 (3.6)	29 (2.7)
Northern Territory	8 (0.5)	5 (0.5)
Australian Capital Territory	31 (2.0)	22 (2.1)
Western Australia	151 (9.8)	98 (9.3)
Highest educational qualification		
No schooling	-	-
Did not complete primary school	6 (0.4)	-
Completed primary school	24 (1.6)	10 (0.9)
Year 10 or equivalent	163 (10.5)	87 (8.2)
Year 11 or equivalent	41 (2.7)	19 (1.8)
Year 12 or equivalent	235 (15.2)	159 (15.0)
A trade, technical certificate or diploma	489 (31.6)	251 (23.7)
A university or college degree	421 (27.2)	372 (35.2)
Postgraduate qualifications	167 (10.8)	159 (15.0)
Work status		
Work full-time	585 (37.8)	566 (53.5)
Work part-time or casual	256 (16.6)	182 (17.2)
Full-time student	27 (1.7)	34 (3.2)
Unemployed and looking for work	69 (4.5)	61 (5.8)
Full-time home duties	102 (6.6)	44 (4.2)
Retired	447 (28.9)	132 (12.5)
Sick or on a disability pension	41 (2.7)	26 (2.5)
Other	19 (1.2)	12 (1.1)
Occupation		
Manager	287 (18.6)	252 (23.8)
Professional	375 (24.3)	259 (24.5)
Technician or trade worker	114 (7.4)	79 (7.5)
Community or personal service worker	90 (5.8)	69 (6.5)
Clerical or administrative worker	268 (17.3)	138 (13.1)
Sales worker	123 (8.0)	104 (9.8)
Machinery operator and driver	58 (3.8)	20 (1.9)
Labourer	139 (9.0)	101 (9.6)
Small business operator	92 (6.0)	35 (3.3)
Marital status		
Single or never married	327 (21.2)	298 (28.2)

Separated or divorced	135 (8.7)	74 (7.0)
Widowed	54 (3.5)	17 (1.6)
Married or living with partner (de facto)	1030 (66.6)	668 (63.2)
Household composition		
Single person	351 (22.7)	249 (23.6)
One parent family with children	77 (5)	74 (7.0)
Couple with children	520 (33.6)	408 (38.6)
Couple with no children	525 (34.0)	266 (25.2)
Group household (i.e. living with two or more people to whom you are NOT related)	73 (4.7)	60 (5.7)
Annual personal income		
$0 to $19,999	244 (15.8)	137 (13.0)
$20,000 to $39,999	402 (26.0)	200 (18.9)
$40,000 to $59,999	233 (15.1)	188 (17.8)
$60,000 to $79,999	238 (15.4)	161 (15.2)
$80,000 to $99,999	160 (10.3)	113 (10.7)
$100,000 to $119,999	102 (6.6)	99 (9.4)
$120,000 to $139,999	60 (3.9)	58 (5.5)
$140,000 to $159,999	39 (2.5)	46 (4.4)
$160,000 to $179,000	24 (1.6)	18 (1.7)
$180,000 or more	44 (2.8)	37 (3.5)
Annual household income		
$0 to $19,999	70 (4.5)	48 (4.5)
$20,000 to $39,999	287 (18.6)	145 (13.7)
$40,000 to $59,999	226 (14.6)	156 (14.8)
$60,000 to $79,999	214 (13.8)	137 (13.0)
$80,000 to $99,999	166 (10.7)	136 (12.9)
$100,000 to $119,999	149 (9.6)	133 (12.6)
$120,000 to $139,999	103 (6.7)	89 (8.4)
$140,000 to $159,999	127 (8.2)	84 (7.9)
$160,000 to $179,000	47 (3.0)	38 (3.6)
$180,000 or more	157 (10.2)	91 (8.6)
Residence		
Capital city and surrounds	1003 (64.9)	766 (72.5)
Regional town with more than 10,000 persons	396 (25.6)	223 (21.1)
A rural or remote location	147 (9.5)	68 (6.4)
PGSI category		
No risk	1193 (77.2)	138 (13.1)
Low risk	246 (15.9)	153 (14.5)
Moderate risk	88 (5.7)	350 (33.1)
Problem gambler	19 (1.2)	416 (39.4)

Analysis

The R statistical programming environment (R Core Team, 2020) was used for all analyses. Tetrachoric correlations were calculated between binary variables using the *psych* package (Revelle & Revelle, 2015). We used two indices of internal consistency to evaluate the candidate items, including coefficient *alpha* based on the average inter-item covariance and coefficient *omega*, which measures the ratio of variance captured by a general factor *g* to that attributable to subfactors (F) via hierarchical confirmatory factor analysis (CFA) (Flora, 2020). Because the items were binary scored, tetrachoric correlations served as the input for all indices of reliability, and logistic links were employed for the CFA. Together, these coefficients speak not only to the reliability of the scale sum, but also to the degree of unidimensionality evinced by the item set.

Further insight into item performance can be gained by applying a unidimensional, two-parameter latent trait model (2P LTM) to the candidate items, also using a logistic link function. This was implemented using the *ltm* package (Rizopoulos, 2006). Like a CFA, this model estimates latent trait scores (Θ), which in the present analysis correspond to the underlying degree of harm experienced. LTM is best thought of as a variant of CFA in a format that emphasises the properties of binary or ordinal response items. Item characteristics are captured via two item parameters: the intercept (α) and slope (β) of the logistic link functions from the latent variable to each binary measure, usually described as the *severity* and *discrimination* parameters in the LTM literature. These parameters can be thought of, respectively, as the degree to which a positive response on an item reflects severe underlying harm (Θ), and how reliably it differentiates differing degrees of underlying harm (Θ). We calculated point-biserial correlations of each item with the latent harm scores, as well as validation outcomes such as the SF-6D.

The *lasso* (least absolute shrinkage and selection operator) is a form of multivariate regression that implements both variable selection and regularisation (Tibshurani, 1996) that addresses some of the limitations of both standard regression and stepwise variable selection in handling collinearity and redundancy in predictors. Rather than simply minimising the mean squared error, it introduces a term that penalises higher magnitude coefficients, thus encouraging a more parsimonious and robust solution. Similar to stepwise regression, the solution will generally include some coefficients thresholded to zero (or excluded from the equation), when they do not contribute sufficient unique information when other variables are included. However, the lasso does not suffer from high dependence on sample variability that is well known to affect stepwise methods. Because of this attractive property, the lasso has been recently applied to psychometric problems in variable/item selection (Gonzalez, 2020; Jacobucci et al., 2019; Sun & Ye, 2019). In this paper the lasso was employed for multivariate evaluation and selection of harm items, predicting the SF-6D and the WHOQOL-BREF subscales as outcomes. The

underlying logic is that if a given harm symptom indexes a meaningful and unique quantum of gambling-related harm, it should be associated with an observable reduction in quality of life as indicated by one or more of these outcomes. Additionally, beta coefficients were constrained to be negative, as a positive coefficient would imply a partial positive effect of a harm on health utility, which is both theoretically undesirable and detrimental to scale performance. This procedure leads to the selection of a subset of harms that demonstrate a robust and unique negative association with each outcome. It is important to note that lasso coefficients are penalized estimates and should not be interpreted as unbiased effect sizes. In this context, the presence of a non-zero coefficient indicates that the item contributed unique predictive information beyond other items, rather than providing a direct estimate of the magnitude of association.

Estimating health impacts and a note on terminology.

To evaluate the final scale, propensity weighted data is used for regression with control for potential confounds to effect of gambling harms on health and wellbeing. We use the term "causal" here in the limited, technical sense common in population health and burden-of-disease research, where regression models are used to approximate counterfactual contrasts between risk factors and outcomes under specified assumptions (e.g., no unmeasured confounding, correct specification). In this tradition, "causal" distinguishes models that aim to isolate direct associations controlling for observed confounders from propensity models (which balance covariates across groups) or purely predictive models. We acknowledge that our data are cross-sectional and cannot establish temporal precedence or rule out reciprocal effects (e.g., negative affect may both result from and predispose to gambling problems). Our usage of "causal" therefore should not be taken as a claim of definitive causal inference, but rather as a shorthand for structural associations interpretable under standard epidemiological assumptions.

Results

Associations between candidate items

Figure 1 displays the matrix of tetrachoric correlations between the 31 candidate harm items. For ease of interpretation, items are algorithmically sorted such that clusters of items with higher mutual correlations are placed together. The magnitude and direction of the correlations were very homogenous, with a mean inter-item correlation of .73, a SD = .08 and IQR [.70, .78]. Accordingly, a three-level colour scaling has been applied to enhance contrast between inter-item correlations. *Reduction of my available spending money, Reduction of my savings,* and *Less spending on recreational expenses such as eating out, going to movies or other entertainment* showed lower correlations with more severe social deviance items such as *Had experiences with violence (including family/domestic violence)*. This was also somewhat the case for the

moderately severe financial harms of *Increased credit card debt* and *Sold personal items*. Similarly, the cluster of *Used my work or study time to gamble, Didn't fully attend to the needs of children,* and *Reduced my contribution to religious or cultural practices*, all indicative of relatively severe social dysfunction, tended to have slightly lower cross-associations overall, but particularly so with milder financial harms. As apparent in the lower left triangle, there were particularly strong mutual associations between emotional impacts. However, it should be emphasised that these patterns were quite subtle differences of degree. The lowest inter-item correlation was .53, which represents a moderately strong positive association. In line with this, alpha reliability for all 31 items was .99 and beta (the minimum split-half reliability) was .97. Similarly, coefficient omega was calculated to be .87, with all items loading strongly (.73-.93) on the general factor, relatively low on subfactors (.21 - .45), and eigenvalues of [*g*: 20.78, F1: 2.3, F2: .80, F3: .96, F4: .46]. As shown in Figure 2, the subfactor loadings mirrored the structure identified above, with some sub-factor structure apparent with respect to emotional affect and depression, financial stressors, and the remaining sub-factor reflecting work, social and cultural dysfunction.

To summarise, (1) the full candidate set of items comprised a very reliable indicator of an underlying general construct of harm, (2) loadings and inter-item associations were very homogenous, (3) some sub-factor structure was apparent, but these facets were subtle and contributed little to the total explained covariance, and (4) no individual items stood out as being incontrovertibly superior or inferior based on simple cross-correlations or factor loadings.

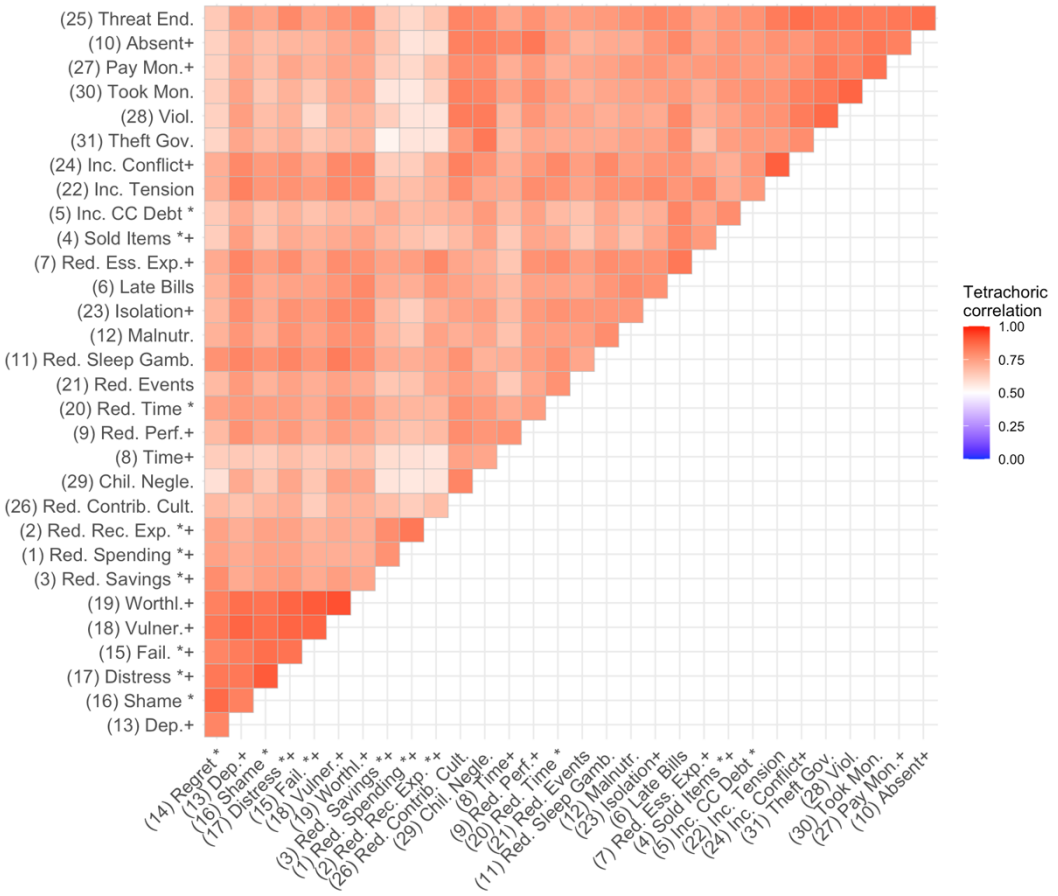

Figure 1. Visualisation of tetrachoric correlations among candidate gambling harm items

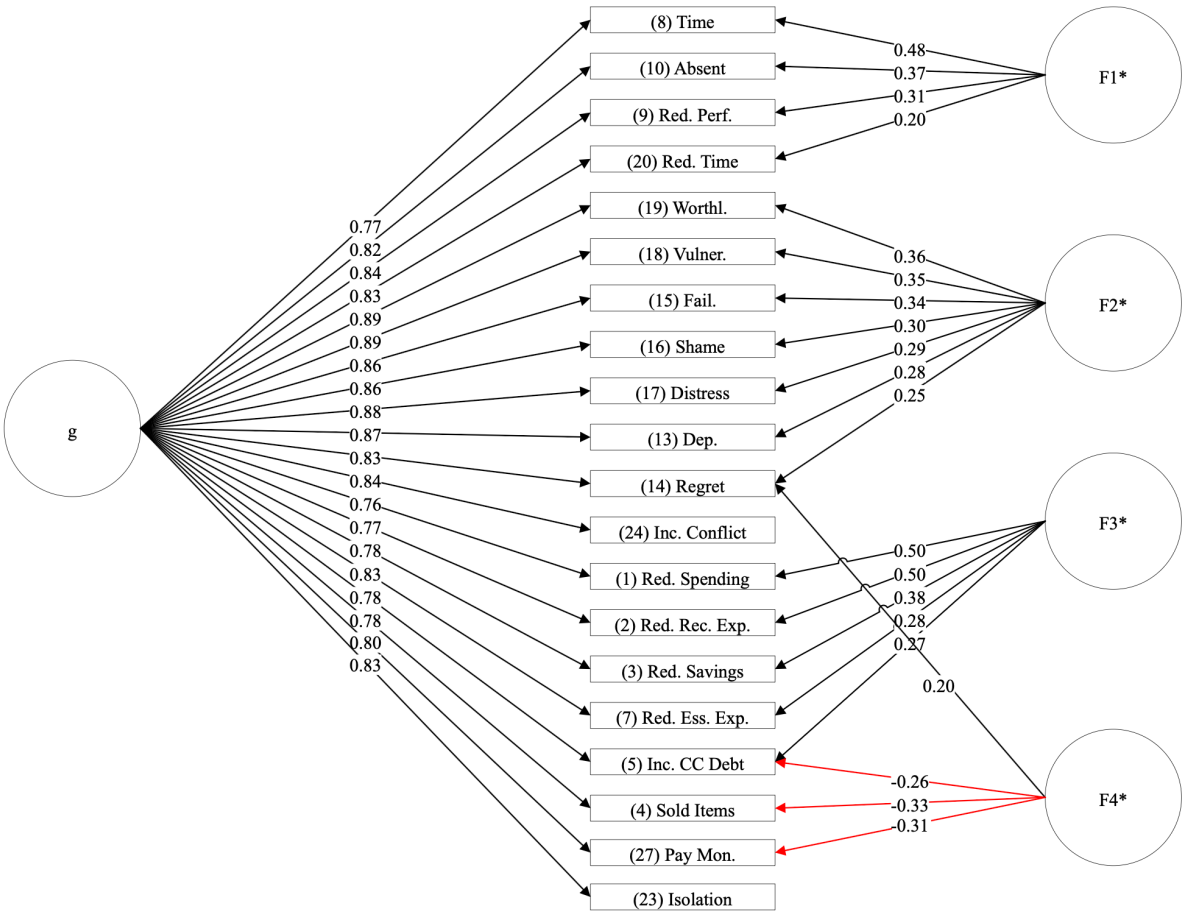

Figure 2. Hierarchical CFA general and subfactor loadings for GHS-20 items

Note: Factors: F1 = Time & preoccupation, F2 = Affect, F3 = Finances, F4 = Work, social and cultural dysfunction.

LTM analysis

The upper section of Table 3 summarises the LTM parameters for each item, along with the point-biserial correlation of each item with the latent factor *r*(Θ), the key validation outcome *r*(SF-6D), as well as the prevalence of positive responses of each item in the present dataset. Item-wise correlations between these parameters are given below. The lower section of Table 3 is a correlation matrix across those indices (column-wise), showing at a glance that items which load more strongly on the harm factor also tend to be more strongly associated with lower health utility, whereas very severe/rare items contribute less to that external benchmark.

Of note is that *r*(Θ) and 2P(β) are very highly correlated (*r* = .91), confirming that both index the degree to which each item indicates the underlying dimension of harm. More interestingly, these indices of internal consistency had a moderately strong correlation with the external measure of health utility (*r*(SF-6D)) of -.65 and -.51 respectively. In other words,

items that had a stronger association with the underlying general factor of harm, also tended to have a stronger association with the external measure of health utility. In contrast, items that had a higher severity (and correspondingly lower prevalence) showed a slightly negative correlation with the external measure: the correlation between *r*(SF-6D) and 2P(α) was -.16. Emotional affect items such as *Vulner.* and *Worthl.* tended to perform best, both in terms of *r*(Θ) and *r*(SF-6D). However, items overall were relatively homogenous according to both criteria. Two further considerations should be kept in mind. First, these two aspects of item performance do not take into account information redundancy between items (**See note 1**). For example, if *Vulner.* and *Worthl.* had a correlation approaching 1, then inclusion of both items in a scale would not be advisable, since they would not provide unique information about either the latent factor or external outcomes. Second, a correlation with either benchmark does not incorporate the prevalence of the item, which affects the information content of the item. For example, given two items showing equivalent bivariate correlational performance in Table 3, if one item had a prevalence of 1% and the second item had a prevalence of 50%, then the more prevalent item would provide far greater information, or (equivalently) explain more variance, in those benchmarks.

Taken together, Table 3 shows that well-discriminating, moderately prevalent items carry more information about observed health decrements than very severe, rare items, consistent with our selection logic.

Table 3. Two-parameter logistic trait (LTM) item statistics and external validity

	r(Θ)	*r*(SF-6D)	2P(α)	2P(β)	N
(1) Red. Spending *+	.37	-.27	1.90	2.23	608
(2) Red. Rec. Exp. *+	.40	-.27	2.27	2.47	554
(3) Red. Savings *+	.41	-.27	1.92	2.36	622
(4) Sold Items *+	.48	-.29	4.40	2.82	229
(5) Inc. CC Debt *	.48	-.24	4.01	2.76	273
(6) Late Bills	.57	-.31	4.82	3.43	266
(7) Red. Ess. Exp.+	.55	-.29	4.62	3.34	278
(8) Time+	.41	-.29	3.39	2.21	281
(9) Red. Perf.+	.53	-.32	5.00	3.38	237
(10) Absent+	.53	-.30	5.64	3.45	178
(11) Red. Sleep Gamb.	.55	-.30	4.15	3.24	329
(12) Malnutr.	.52	-.28	4.17	3.07	298
(13) Dep.+	.59	-.36	4.55	3.67	337
(14) Regret *	.44	-.29	2.35	2.55	546
(15) Fail. *+	.51	-.32	3.61	3.06	385
(16) Shame *	.50	-.29	3.26	3.05	442
(17) Distress *+	.58	-.31	4.11	3.59	386

(18) Vulner.+	.62	-.35	4.89	3.83	317
(19) Worthl.+	.61	-.35	4.65	3.72	331
(20) Red. Time *	.50	-.31	3.60	2.91	361
(21) Red. Events	.48	-.26	3.60	2.66	321
(22) Inc. Tension	.59	-.31	4.90	3.65	289
(23) Isolation+	.54	-.33	4.16	3.21	322
(24) Inc. Conflict+	.59	-.31	5.17	3.74	269
(25) Threat End.	.59	-.30	5.82	3.96	225
(26) Red. Contrib. Cult.	.51	-.27	4.62	3.04	233
(27) Pay Mon.+	.53	-.31	5.03	3.35	230
(28) Viol.	.52	-.28	5.86	3.56	170
(29) Chil. Negle.	.50	-.27	5.05	3.12	195
(30) Took Mon.	.52	-.26	6.05	3.62	160
(31) Theft Gov.	.49	-.27	6.37	3.68	141

	Correlations				
	***r*(Θ)**	***r*(SF-6D)**	**2P(α)**	**2P(β)**	**N**
r(Θ)	-	-.65	.67	.91	-.51
r(SF-6D)		-	-.16	-.51	.00
2P(α)			-	.84	-.93
2P(β)				-	-.61

Note: *The upper panel lists, for each candidate harm item: point-biserial correlation with the latent harm factor r(θ), correlation with health utility r(SF-6D), LTM severity P(α), LTM discrimination P(β), and item prevalence (N positive responses). The lower panel reports the correlations among these column-wise indices across all items (i.e., how strongly items that discriminate well also correlate with SF-6D, how severity relates to SF-6D, etc.).*

Multivariate item selection via the lasso

Regularised multivariate regressions were conducted on the SF-6D and sub-scales of the WHOQOL-BREF using lasso, and constraining coefficients to be zero or less. Beta coefficient estimates are summarised in Table 4, ranked by aggregate unique effect magnitude. The reader is reminded that these reflect partial effects of a set of items with a strong unidimensional structure and accordingly, significant collinearity. Thus, a zero reported coefficient does not imply no relationship of the harm item with the SF-6D, only that it does not meet the lasso threshold for contribution of *additional* explanatory information, given the potential inclusion of all other harm items. Items for the GHS-20 (item wordings provided in Appendix 1) were selected based on the criteria of providing unique explanatory information on the SF-6D or WHOQOL-BREF subscales. Sixteen of the thirty-one coefficients were non-zero, reflecting robust contribution of independent information in the outcome, in the context of all other items. Six of the ten items comprising the GHS-10 were re-identified by this method. Emo/Psych items and depression (Health) showed the strongest effects. Financial items, including milder items, were

also strongly represented in the list of non-zero coefficient estimates. However, items were identified from each of the six harm domains.

Table 4. Penalised lasso coefficients for simultaneous regressions predicting health and quality of life outcomes on candidate harms

Category	Harms Item	Outcome				
		SF-6D	WHOQOL-BREF			
			Phys.	Psych.	Relations.	Environ.
Health	(13) Dep.+	-0.26	-0.18	-0.13	-0.12	-0.08
Emo/Psych	(19) Worthl.+	-0.04	-0.08	-0.14	-0.10	-0.18
Emo/Psych	(15) Fail.*+	-0.02	-0.12	-0.12	-0.12	-0.07
Emo/Psych	(17) Distress*+	0.00	0.00	-0.11	-0.13	-0.16
Relationships	(23) Isolation+	-0.14	-0.12	0.00	0.00	0.00
Emo/Psych	(18) Vulner.+	-0.14	-0.04	0.00	0.00	0.00
Work/Study	(8) Time+	-0.15	0.00	0.00	0.00	0.00
Financial	(7) Red. Ess. Exp.+	0.00	0.00	0.00	0.00	-0.14
Financial	(2) Red. Rec. Exp.*+	0.00	-0.04	-0.03	0.00	-0.05
Work/Study	(9) Red. Perf.+	-0.10	0.00	0.00	0.00	0.00
Financial	(1) Red. Spending*+	-0.03	0.00	-0.04	0.00	0.00
Financial	(3) Red. Savings*+	0.00	-0.06	0.00	0.00	0.00
Work/Study	(10) Absent+	-0.05	0.00	0.00	0.00	0.00
Relationships	(24) Inc. Conflict+	-0.03	0.00	0.00	0.00	0.00
Financial	(4) Sold Items*+	-0.02	0.00	0.00	0.00	0.00
Soc. Deviance	(27) Pay Mon.+	-0.01	0.00	0.00	0.00	0.00
Relationships	(20) Red. Time*	0.00	0.00	0.00	0.00	0.00
Financial	(5) Inc. CC Debt*	0.00	0.00	0.00	0.00	0.00
Financial	(6) Late Bills	0.00	0.00	0.00	0.00	0.00
Health	(11) Red. Sleep Gamb.	0.00	0.00	0.00	0.00	0.00
Health	(12) Malnutr.	0.00	0.00	0.00	0.00	0.00
Emo/Psych	(14) Regret*	0.00	0.00	0.00	0.00	0.00
Emo/Psych	(16) Shame*	0.00	0.00	0.00	0.00	0.00
Relationships	(21) Red. Events	0.00	0.00	0.00	0.00	0.00
Relationships	(22) Inc. Tension	0.00	0.00	0.00	0.00	0.00
Relationships	(25) Threat End.	0.00	0.00	0.00	0.00	0.00
Soc. Deviance	(26) Red. Contrib. Cult.	0.00	0.00	0.00	0.00	0.00
Soc. Deviance	(28) Viol.	0.00	0.00	0.00	0.00	0.00
Soc. Deviance	(29) Chil. Negle.	0.00	0.00	0.00	0.00	0.00
Soc. Deviance	(30) Took Mon.	0.00	0.00	0.00	0.00	0.00
Soc. Deviance	(31) Theft Gov.	0.00	0.00	0.00	0.00	0.00

Note: SF-6D = Short Form 6-Dimension; WHOQOL-BREF = World Health Organization Quality of Life-BREF. Non-zero penalised coefficients shaded, * item included in the original GHS-10, + item identified in the present analysis.

Validity and internal consistency of the GHS-20

The union of 10 GHS-10 items and lasso-identified items comprised 20 harms that comprise the GHS-20 (text of items provided in Appendix 1). The sum of the additional 10 items, excluding common items with the GHS-10 is labelled GHS-20^{-10} and was also analysed to provide a form of split-half reliability. Table 5 summarises bivariate associations of the harm scales and the PGSI with the two major health and quality of life validation measures. As shown in Table 5, the two 10-item halves of the GHS-20 were moderately strongly correlated at r = .78. The GHS-20^{-10} subset showed a slightly stronger correlation (-.41) with the SF-6D than GHS-20, the GHS-10 or the PGSI (-.37). However, the GHS-10 and GHS-20 had stronger associations with the PGSI (.77, .78 respectively) than the subset of new items (.70). Associations of all measures with the SF-6D tended to be higher than with the WHOQOL-BREF. Alpha and omega coefficients for the GHS-20 were .98 and .90, respectively. Figure 2 illustrates a 4-subfacet hierarchical CFA on the GHS-20 items that corresponds to the calculated omega. Loadings on the general factor are strong and homogenous and sub-facet variance is small with eigenvalues of {G: 13.66, F1: .62, F2: .77, F3: .89, F4: .50}. Nevertheless, coherent clustering can be seen with respect to item content: time and preoccupation (F1), affect (F2) and finances (F3). F4 explains marginal variance but appears to capture a small negative residual association between propensity to feel regret, and some indices of 'irresponsible' financial behaviour. Confirmatory factor analysis demonstrated excellent structural validity for the GHS-20, with the unidimensional model showing strong fit (CFI = .996, TLI = .996, RMSEA = .041, SRMR = .052) and the hierarchical four-factor model showing even better fit (CFI = .999, TLI = .998, RMSEA = .026, SRMR = .038). These indices confirm that while the scale functions well as a unidimensional measure suitable for sum scoring, it also accommodates meaningful subfactor structure as identified in the omega analysis.

To summarise, coefficient omega, CFA eigenvalues and factor loadings confirm that the GHS-20 demonstrated strong internal consistency and unidimensionality. Bivariate negative correlations with health and life quality validation measures were not markedly higher than the GHS-10, and both the GHS-10 and the GHS-20 showed almost identical correlations with the PGSI.

Table 5. Bivariate correlations of harm scales with the PGSI, and health and quality of life measures

	GHS-20	GHS-20^{-10}	GHS-10	PGSI	SF-6D
GHS-20^{-10}	.84				
GHS-10	.98	.78			
PGSI	.78	.70	.77		
SF-6D	-.38	-.41	-.36	-.37	
WHOQOL-BREF	-.31	-.31	-.30	-.30	.63

Note: N = 2,602, All correlations significant $p < .01$. GHS-20 = 20-item Gambling Harms Scale; GHS-10 = 10-item Gambling Harms Scale (formerly SGHS); GHS-20^{-10} = the 10 additional items in the GHS-20 not present in the GHS-10; PGSI = Problem Gambling Severity Index; SF-6D = Short Form 6-Dimension health utility measure; WHOQOL-BREF = World Health Organization Quality of Life-BREF. Associations between the GHS-20-10 and the GHS-10 can be thought of as a form of split-half reliability.

Benchmarking the GHS-20 to health utility

The final stage in the analysis was to calculate the magnitude of the decline in health utility, conditional on GHS-20 scores. We applied a propensity weighting and control framework (Browne et al., 2020), and the same methodology that has been used before (Browne et al., 2022) . A total of 1,102 gamblers who scored zero on the GHS-20 formed the reference group. The first column of Table 6 describes a logistic model of the propensity of participants to be a harmed case, i.e. score 1+ on the GHS-20 (N = 1,501) as a function of demographic characteristics. The predicted probabilities are used to calculate case weights in order to balance the influence of these factors for subsequent causal models (columns 2 and 3 of Table 6). These models also control for known comorbidities (e.g., smoking, disability) of gambling-related harm that may be confounded with health utility. Selection of covariates was guided by a prior systematic framework for estimating the indirect health impacts of gambling, which identified comorbidities commonly associated with gambling problems (Browne et al., 2020). In the present dataset, these were operationalised as alcohol use (AUDIT-C), disability/sickness pension status, education, and income, along with demographic factors (Table 6). This approach ensured that the estimated associations between gambling harms and health utility were not confounded by key comorbid conditions.

The first causal model shows mean decrements associated with arbitrary cut-offs of the GHS-20: 0 versus: [1-2], [3-5], [6-10], [11-20], and the second causal model incorporates only a single linear term for the raw GHS-20 score. We also calculated a Generalised Additive Model (GAM) for causal effects, which makes no assumptions regarding group cut-offs or linearity, and allows for fitting of any smooth relationship between the GHS-20 and the SF-6D. Interestingly, plotting the conditional effects of this model showed a nearly linear relationship. Separating the linear and smooth terms showed a significant effect for curvilinearity, $F(\text{edf} = 1.631) = 1.137$,

p = .002, but this improved the explained variance only marginally: 31.1% to 31.4%. Figure 3 shows the GAM estimated decrements by GHS-20 score, with a reference line to indicate linearity. Some concavity is evident in the marginal SF-6D scores, indicating some degree of saturation. This is to be expected, as progressively more symptoms are reported, there is a logical bound to the degree further symptoms indicate further impact to health. The utility decrements for the GHS-10 (formerly SGHS) was done using an identical method and dataset as for the GHS-20, as described here, but using scores on the 10-item version, rather than the 20-item version, as the instrumental variable.

Table 6. Model summaries and beta coefficients for propensity and causal models

		Beta coefficients (SE)		
Model		**Propensity**	**Causal (categorised)**	**Causal (score)**
Regression type		**Logistic**	**OLS**	**OLS**
DV		***P*(GHS20>0)**	**SF-6D**	**SF-6D**
Intercept		2.008** (0.252)	0.840** (0.009)	0.835** (0.008)
GHS-20 categories	0		-	
	1-2		-0.018** (0.006)	
	3-5		-0.043** (0.008)	
	6-10		-0.083** (0.008)	
	11-20		-0.129** (0.008)	
GHS-20 sum				-0.012** (0.001)
Alcohol consumption	Non-drinker		-	-
(AUDIT-C)	Non-risky		0.006 (0.008)	0.006 (0.008)
	Risky		0.005 (0.007)	0.006 (0.007)
Age	Linear	-0.035** (0.003)	0.169 (0.123)	0.142 (0.122)
	Quadratic		-0.314** (0.117)	-0.318** (0.116)
	Cubic		-0.202 (0.116)	-0.188 (0.115)
Gender	Male	-	-	-
	Female	-0.188* (0.088)	-0.017** (0.005)	-0.017** (0.005)
Country of birth	Overseas	-	-	-
	Australia	-0.184 (0.106)	-0.015** (0.006)	-0.015** (0.006)

Education	Secondary or less	-		
	Trade/Cert	-0.041 (0.113)		
	Tertiary	0.192 (0.118)		
	Postgrad	0.401** (0.153)		
Personal income		0.101** (0.029)		
Household income		-0.116** (0.024)		
Mother's highest education		-0.066** (0.019)		
Sick /disability pension	No		-	-
	Yes		-0.128** (0.015)	-0.129** (0.015)
Recreational drug use	No		-	-
	Yes		-0.017* (0.008)	-0.016* (0.008)
Cigarettes per day	0		-	-
	< 10		-0.015* (0.007)	-0.015* (0.007)
	10+		-0.012 (0.007)	-0.012 (0.007)
Past year diagnosis of psychological disorder…	Mood		-0.052** (0.009)	-0.051** (0.009)
	Anxiety		-0.064** (0.007)	-0.062** (0.007)
	Personality		-0.011 (0.013)	-0.006 (0.013)
	Any other		-0.031** (0.012)	-0.026* (0.012)
	N	2,603	2,603	2,603
	R^2		0.312	0.314
	Adjusted R^2		0.307	0.310
	Log Likelihood	-1,655.800		
	Akaike Inf. Crit.	3,331.601		
	Residual Std. Error		0.163 (df = 2583)	0.163 (df = 2586)
	F Statistic		61.724** (df = 19; 2583)	74.046** (df = 16; 2586)

Note:*p<0.05; **p<0.01. GHS-20 = 20-item Gambling Harms Scale; AUDIT-C = Alcohol Use Disorders Identification Test-Consumption; OLS = Ordinary Least Squares; SF-6D = Short Form 6-Dimension health utility measure. The propensity model uses logistic regression to predict the probability of being a harmed case (GHS-20 > 0). The causal models use propensity score weighting to balance demographic factors and control for comorbidities when estimating the relationship between GHS-20 scores and health utility

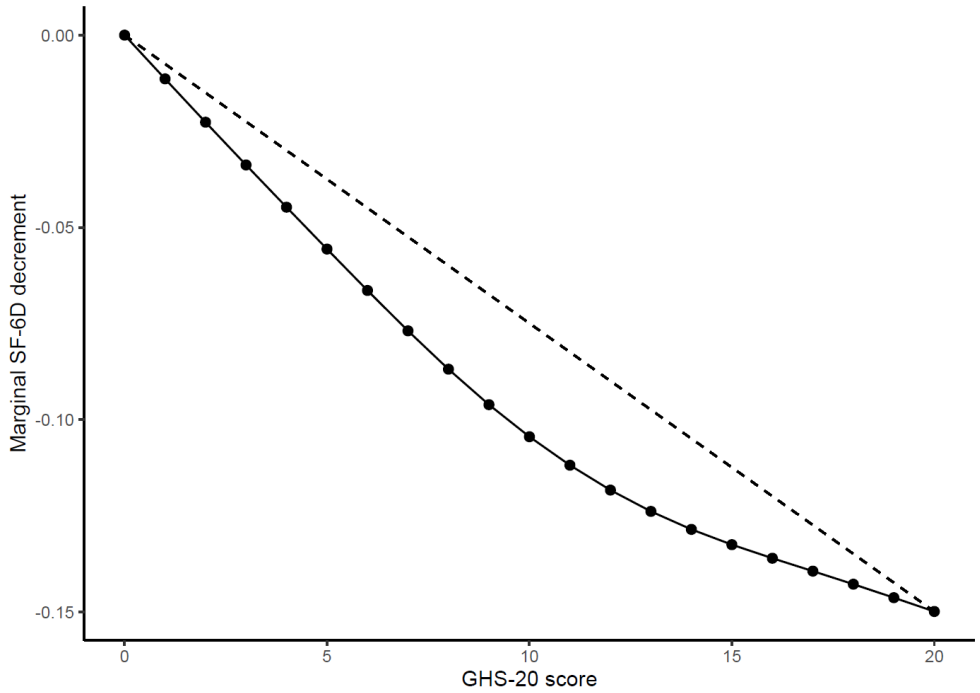

Figure 3. SF-6D decrements conditional on GHS-20 score

Note: Dashed reference line indicates departure from linearity. SF-6D = Short Form 6-Dimension health utility measure; GHS-20 = 20-item Gambling Harms Scale. Dashed reference line indicates departure from linearity.

Discussion

In Hamlet, Claudius laments that "When sorrows come, they come not single spies, but in battalions". This truism applies to the primary finding of the present analysis, which is that negative incidents and symptoms arising from gambling do not happen in isolation, but rather reflect an underlying malaise with significant implications for health and wellbeing. In short, the reliability and validity of the construct of gambling related harm is well supported. That is, each instance of harm correlates well with the whole, and the totality of harm is predictive of decrements to health utility. The strong structural connections between specific harm symptomology make the task of the psychometrician easier, in that a sufficiently large set of symptoms, when aggregated, can be almost guaranteed to do relatively well at capturing the underlying degree of impact to the gambler. This is illustrated by the homogenous positive intercorrelations between individual harms and loadings on a common factor, and most simply by the +.78 correlation between the 10 additional items for the GHS-20 and the original GHS-10. Although some items showed weak or inconsistent loadings on subfactors in the hierarchical CFA, all 20 items loaded strongly and homogeneously on the general factor. This supports the interpretation of the GHS-20 as a primarily unidimensional measure of gambling harm. Subfactor structures were examined for theoretical interest and potential future applications, but they contributed small-to-moderate unique variance beyond the general harm factor.

However, a corollary of strong unidimensional structure of gambling harm symptomatology is that there is a law of decreasing returns when extending the original scale. Although the GHS-20 was designed to maximise criterion validity with respect to health and wellbeing—i.e., items were selected for their explanatory association with decrements in health utility and quality-of-life outcomes, not to optimise out-of-sample prediction (see Shmueli, 2010), it provides either marginal or no improvement to the GHS-10 in terms of predicting outcomes. Nevertheless, as predicted by theory, like the GHS-10, the GHS-20 is associated (r = -.38) with the key criterion measure, the SF-6D. Researchers are advised to employ the GHS-20 when time/space permits, when more precision in marginal health utility impact is required, or better coverage of the domains of harm is desired. In statistical terms, the GHS-10 is a perfectly feasible substitute when this is not the case, being correlated with the GHS-20 at .98, and having similar relationships with key outcomes.

A strength of the present development effort is the use of propensity-weighted multivariate causal modelling: i.e. balancing harmed and unharmed individuals with respect to demographics and controlling for major comorbidities. Linear and non-linear fits show that increasing GHS-20 scores are associated with unique and progressive decrements to health utility. Scores as low as 1-2 show a statistically significant decrement to health. The scoring tables provided in Appendices 2 & 3 of Browne et al. (2023) are commended for use in population surveys to calculate an aggregate 'burden of harm' for gamblers. We intend to publish similar instruments and scoring instructions for matching instruments for those affected by another's gambling.

The item content and sub-factor structure of the GHS-20 deserve special mention. The additional selected items yield a scale that includes items from all domains of harm as conceptualised by Langham et al. (2015). As illustrated in Figure 2, and captured by diagnostics such as coefficient omega, although the scale is highly unidimensional, there is some detectable sub-factor structure. Thematic clustering can be seen with respect to reallocation of time from other life commitments, financial impacts, and psychological distress. The first two clusters map onto the two theoretical mechanisms by which harm from gambling is thought to occur, which is the excessive attrition of time and money from the gambler. Other items such as social isolation and increased conflict in relationships, along with psychological distress, represent more direct impacts on health and wellbeing. Nevertheless, theoretical structure does not necessarily match with statistical structure: a person experiencing gambling-related harm is likely to experience symptomatology across all domains. Whilst we would not discourage research on specific harm domains, all psychometric evidence on the GHS-10 (formerly SGHS) and GHS-20 points to a unitary structure.

A note of caution should be sounded regarding the interpretation of raw correlations reported for descriptive purposes in this study. Harms from

gambling vary markedly in their prevalence, with lower prevalence tending to be indicative of more severe gambling problems and greater impact to health utility. To illustrate, given two binary items with identical and equal correlations with an outcome of interest, if item (A) has a prevalence of 1 in 1000 and (B) has a prevalence of 1 in 10, then (B) would explain approximately 100 times more variance in a population-representative sample, and is far more desirable from a psychometric perspective. In short, all else being equal, more prevalent items provide more information about an outcome. The lasso regression item-selection approach taken in this paper, which also considers inter-item redundancy (i.e. common covariation with the outcome), is therefore the preferred benchmark for item evaluation.

Applications and clinical implications

While the GHS-20 was primarily developed for population surveillance and research purposes, it offers several advantages for clinical practice in Australia's gambling treatment settings. The scale's inclusion of moderately prevalent symptoms across all harm domains provides clinicians with a comprehensive yet efficient screening tool that captures the multifaceted nature of gambling-related harm. Unlike problem gambling screens that focus on addiction criteria, the GHS-20's emphasis on experienced harms aligns with person-centred approaches increasingly adopted in Australian mental health services.

The scale's benchmarking to health utility provides clinicians with an evidence-based framework for communicating impact to clients and other health professionals. For instance, a GHS-20 score of 5 corresponds to a measurable health decrement comparable to moderately-severe chronic conditions, offering a concrete reference point for treatment planning. The item content itself can facilitate clinical conversations, as each endorsed item represents a specific harm domain that warrants exploration. This is particularly valuable in Australia's integrated care settings where gambling harm often co-occurs with other mental health and substance use concerns.

Furthermore, the GHS-20's sensitivity to low-level harms (detecting significant decrements at scores of 1-2) supports early intervention approaches prioritised in Australian public health policy. This allows services to identify and support individuals experiencing gambling harm before problems escalate, potentially reducing the trajectory toward more severe gambling disorders.

Limitations and Future Directions

Several limitations warrant consideration. First, the sample was drawn from a commercial panel during COVID-19 restrictions, potentially affecting representativeness despite quota sampling. Victorian residents were excluded due to lockdowns limiting venue-based gambling access, which may impact generalizability given Victoria's high EGM density. Second, the cross-sectional design precludes causal inference about the relationship between gambling harms and health utility, though propensity

weighting strengthens these associations. Third, while the scale demonstrates excellent psychometric properties, test-retest reliability and sensitivity to change remain unexamined, limiting conclusions about its utility for monitoring individual progress over time. Fourth, the binary response format, while maintaining consistency with the original GHS-10, may reduce sensitivity compared to graded response options. Also, validation was conducted exclusively within an Australian context; cross-cultural validity requires investigation given potential cultural differences in gambling harm conceptualization and expression. Tetrachoric correlations can be less stable with low base rates. While most items had adequate endorsement, several items in the final GHS-20 had relatively low prevalence (7-10%). However, our conclusions regarding dimensionality are supported by multiple complementary analytic approaches beyond the correlation matrix.

Our use of a bifactor model to examine factor structure was driven by our model of an essentially unidimensional nature of gambling harms. Alternative factor analytic approaches, such as second-order factor models, might provide different insights into the relationship between general and specific harm dimensions. While the bifactor model was well-suited to our goal of confirming unidimensionality while checking for residual subfactor variance, future research employing second-order models could potentially identify more distinct harm clusters or validate domain-specific measurement. Such alternative approaches may be particularly valuable for researchers interested in profiling different patterns of harm or developing domain-specific interventions, rather than our focus on overall harm severity. Future research could also examine whether more differentiated subfactor structures emerge when analyses are restricted to higher-severity cases (e.g., PGSI moderate risk and problem gamblers).

Conclusion

The principal contribution of this work is to provide a new scale for gambling-related harm, the GHS-20, that covers all dimensions of harm identified by Langham et al. (2015), while also providing a metric for how scores on the scale imply a decrement to a gambler's health and wellbeing. Having a wide representation of the different types of harm will give end-users confidence in employing a comprehensive measure that is well-related to meaningful health outcomes and the PGSI. Having a measure of harm, like the GHS-20, that is related to gambling problems but is conceptually independent of gambling as a mental health condition provides unique advantages. Measuring harm independently of mental illness allows for a quantification of the harm experienced by people with different levels of comorbidities, but also provides for a recognition that it is at least possible to experience gambling-harm without having a mental health condition. This brings gambling studies into line with other health-related fields, which recognise that people can experience harm from (e.g.) alcohol and drug use without necessarily having a substance dependency. Such an

endeavour is not 'lowering the bar' for gambling problems, but rather recognising that health-related impacts can impact, to varying degrees, a broader class of gamblers.

Note 1: Excessive correlations between items can certainly be taken into account within an omnibus CFA / LTM framework for item selection (Browne et al., 2018). However, in the interest of accessibility, we opted for the use of more well-established methods.

Appendix 1 Item content and scoring instructions for the GHS-20

Presentation instructions

• The information below includes the wording, item content, and presentation of the GHS-20.
• Where possible the presentation order of the 20 items should be randomised

During the last 12 months, did any of these issues occur as a result of **your gambling**?

Item	No	Yes
Reduction of my available spending money		
Less spending on recreational expenses such as eating out, going to movies or other entertainment		
Reduction of my savings		
Sold personal items		
Increased credit card debt		
Less spending on essential expenses such as medications, healthcare and food		
Used my work or study time to gamble		
Reduced performance at work or study (i.e. due to tiredness or distraction)		
Was absent from work or study		
Increased experience of depression		
Had regrets that made me feel sorry about my gambling		
Felt like a failure		
Felt ashamed of my gambling		
Felt distressed about my gambling		
Felt insecure or vulnerable		
Felt worthless		
Spent less time with people I care about		
Social isolation (felt excluded or shut-off from others)		
Experienced greater conflict in my relationships (arguing, fighting, ultimatums)		
Promised to pay back money without genuinely intending to do so		

Scoring instructions

• A 'no' response to an item should be coded as 0 and a 'yes' response coded as 1.
• Sum scores for all of the 20 items; total scores should range from 0 to 20
• Higher scores indicate higher levels of gambling-harm

STATEMENT OF COMPETING INTERESTS

None

ETHICS APPROVAL

Ethical approval for this study was received from the Central Queensland University Ethics Committee (approval #22341) on 11/05/2020.

RELATIVE CONTRIBUTIONS

All authors conceived of the study. MB conducted the analyses and wrote the first draft of the paper. PN prepared the draft for journal submission. VR managed operations for data collection and screening. AR and MR checked statistical analyses. NH and EL edited the final version of the MS. All authors approved of the final version.

ACKNOWLEDGEMENTS AND FUNDING SOURCES

MB led a grant from the Victorian Responsible Gambling Foundation that funded this project.

RESEARCH PROMOTION

This study was conducted to develop and validate the 20-item Gambling Harms Scale (GHS-20), a new measure designed to provide comprehensive coverage across all six domains of gambling harm. The scale was benchmarked against health utility (SF-6D) using propensity score weighting to create a robust tool for public health assessment. The findings demonstrate that the GHS-20 is a highly reliable measure and that harm scores have a significant, near-linear negative relationship with health utility, with statistically significant health decrements observed even at low scores.

References

Australian Institute of Family Studies. (2025). *National Gambling Prevalence Study Pilot – Key findings*. Australian Gambling Research Centre, AIFS.

Barbosa, C., Bray, J. W., Dowd, W. N., Barnosky, A., & Wittenberg, E. (2021). SF-6D utility scores for alcohol use disorder status and alcohol consumption risk levels in the US population. *Addiction*, *116*(5), 1034-1042. doi:10.1111/add.15224

Brazier, J., & Deverill, M. (1999). A checklist for judging preference-based measures of health-related quality of life: learning from psychometrics. *Health Economics, 8*(1), 41–51. doi: 10.1002/(sici)1099-1050(199902)8:1<41::aid-hec395>3.0.co;2-#

Brazier, J., Ratcliffe, J., Salomon, J. A., & Tsuchiya, A. (2017). Measuring and valuing health benefits for economic evaluation. Oxford University Press. doi: 10.1093/med/9780198725923.001.0001

Brazier, J., Roberts, J., Tsuchiya, A., & Busschbach, J. (2004). A comparison of the EQ-5D and SF-6D across seven patient groups. *Health Economics, 3*(9), 873–884. doi:10.1002/hec.866

Browne, M. (2019). Measuring harm from gambling and estimating its distribution in the population. In H. B. Jones, C. Dickson, C. Dunand, & O. Simon (Eds), *Harm reduction for problem gambling: A public health approach*. Routledge. doi: 10.4324/9780429490750-3

Browne, M., Goodwin, B., & Rockloff, M. (2017). Validation of the Short Gambling Harm Screen (SGHS): A tool for assessment of harms from gambling. *Journal of Gambling Studies, 34*(2), 499-512. doi:10.1007/s10899-017-9698-y

Browne, M., Newall, P., Rawat, V., Tulloch, C., Rockloff, M., Li, E., Hing, N., Russell, A. M. T., & Begg, S. (2023). *The Gambling Harms Scales: Instruments to assess impact to gamblers and affected others that are benchmarked to health utility*. Victorian Responsible Gambling Foundation.

Browne, M., Rawat, V., Newall, P., Begg, S., Rockloff, M., & Hing, N. (2020). A framework for indirect elicitation of the public health impact of gambling problems. *BMC Public Health, 20*, 1-14. doi: 10.1186/s12889-020-09813-z

Browne, M., Rawat, V., Tulloch, C., Murray-Boyle, C., & Rockloff, M. (2021). The evolution of gambling-related harm measurement: Lessons from the last decade. International *Journal of Environmental Research and Public Health, 18*(9), 4395. doi: 10.3390/ijerph18094395

Browne, M., & Rockloff, M. J. (2018). Prevalence of gambling-related harm provides evidence for the prevention paradox. *Journal of Behavioral Addictions, 7*(2), 410-422. doi:10.1556/2006.7.2018.41

Browne, M., Rockloff, M., & Rawat, V. (2018). An SEM algorithm for scale reduction incorporating evaluation of multiple psychometric criteria. *Sociological Methods & Research, 47*(4), 812-836. doi:10.1177/0049124116661580

Browne, M., Russell, A. M. T., Begg, S., Rockloff, M. J., Li, E., Rawat, V., & Hing, N. (2022). Benchmarking gambling screens to health-state utility: the PGSI and the SGHS estimate similar levels of population gambling-harm. BMC Public Health, 22(1), 839. doi: 10.1186/s12889-022-13243-4

Browne, M., Volberg, R., Rockloff, M., & Salonen, A. H. (2020). The prevention paradox applies to some but not all gambling harms: Results from a Finnish population-representative survey. *Journal of Behavioral Addictions, 9*(2), 371-382. doi: 10.1556/2006.2020.00018

Burdine., J. N., Felix, M. R., Abel, A. L., Wiltraut, C. J., & Musselman, Y. J. (2000). The SF-12 as a population health measure: An exploratory examination of potential for application. *Health Services Research*, *35*(4), 885–904.

Busija, L., Pausenberger, E., Haines, T. P., Haymes, S., Buchbinder, R., & Osborne, R. H. (2011). Adult measures of general health and health-related quality of life: medical outcomes study short form 36-item (SF-36) and short form 12-item (SF-12) health surveys, nottingham health profile (NHP), sickness impact profile (SIP), medical outcomes study short form 6D (SF-6D), health utilities index mark 3 (HUI3), quality of well-being scale (QWB), and assessment of quality of life (AQoL). *Arthritis Care & Research, 63*(S11), S383-S412. doi: 10.1002/acr.20541

Delfabbro, P., & King, D. (2017). Prevention paradox logic and problem gambling: Does low-risk gambling impose a greater burden of harm than high-risk gambling? *Journal of Behavioral Addictions, 6*(2), 163–167. doi:10.1556/2006.6.2017.022

Ferris, J. A., & Wynne, H. J. (2001). *The Canadian problem gambling index: Final report.* Ottowa, ON: Canadian Centre on Substance Abuse. doi: 10.5683/SP3/JBMMY0

Flora, D. B. (2020). Your coefficient alpha is probably wrong, but which coefficient omega is right? A tutorial on using R to obtain better reliability estimates. *Advances in Methods and Practices in Psychological Science, 3*(4), 484-501. doi:10.1177/2515245920951747

Gooding, N. B., Allami, Y., Delfabbro, P., Williams, R. J., Parke, J., Volberg, R. A., & Hodgins, D. C. (2024). Validation of the gambling harm measure across three independent samples. *International Gambling Studies*. doi: 10.1080/14459795.2024.2432318

Gonzalez, O. (2020). Psychometric and machine learning approaches to reduce the length of scales. *Multivariate Behavioral Research*, *56*(6), 1-17. doi:10.1080/00273171.2020.1781585

Hunger, M., Baumert, J., Holle, R. (2011). Analysis of SF-6D index data: Is beta regression appropriate? *Value Health, 14,* 759-767. doi:10.1016/j.jval.2010.12.009

Jacobucci, R., Brandmaier, A. M., & Kievit, R. A. (2019). A practical guide to variable selection in structural equation modeling by using regularized multiple-indicators, multiple-causes models. *Advances in Methods and Practices in Psychological Science, 2*(1), 55-76. doi:10.1177/2515245919826527

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), 1-23. doi:10.1186/s12889-016-2747-0

McLauchlan, J., Browne, M., Russell, A. M., & Rockloff, M. (2020). Evaluating the reliability and validity of the Short Gambling Harm Screen: Are binary scales worse than Likert scales at capturing gambling harm? *Journal of Gambling Issues, 44*. doi:10.4309/jgi.2020.44.6

Mou, C., Ferguson, E., Tunney, R. J., & James, R. J. (2025). Assessing the factor structure of the Problem and Pathological Gambling Measure (PPGM). *International Gambling Studies*. doi: 10.1080/14459795.2025.2481846

Murray-Boyle, C., Browne, M., Rockloff, M.R. & Flenady, T. (2021). Opportunity costs or not? Validating the Short Gambling Harm Screen against a set of "unimpeachable" negative impacts. *Journal of Clinical Medicine, 10*(3), 549. doi:10.3390/jcm10030549

Queensland Government Statistician's Office, Queensland Treasury. (2025). *Australian gambling statistics (40th edition, 1998–99 to 2023–24)*.

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Revelle, W., & Revelle, M. W. (2015). Package 'psych'. *The comprehensive R archive network, 337*, 338.

Rizopoulos, D. (2006) ltm: An R package for latent variable modelling and item response theory analyses. *Journal of Statistical Software, 17*(5), 1–25. doi:10.18637/jss.v017.i05

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*(1), 41–55. doi: 10.1093/biomet/70.1.41

Shmueli, G. (2010). To Explain or to Predict? *Statistical Science, 25*(3), 289–310. DOI: 10.1214/10-STS330.

Sun, J., & Ye, Z. (2019). A LASSO-based method for detecting item-trait patterns of replenished items in multidimensional computerized adaptive testing. *Frontiers in Psychology, 10*, 1944. doi:10.3389/fpsyg.2019.01944

Tibshirani, R. (1996). Regression Shrinkage and Selection via the lasso. *Journal of the Royal Statistical Society, 58*(1), 267–88. doi:10.1111/j.2517-6161.1996.tb02080.x

Wardle, H., Degenhardt, L., Marionneau, V., Reith, G., Livingstone, C., Sparrow, M., ... & Saxena, S. (2024). The Lancet public health commission on gambling. *The Lancet Public Health, 9*(11), e950-e994. doi: 0.1016/S2468-2667(24)00167-1

Wardle H, Reith G, Best D, McDaid D, & Platt S. (2018). *Measuring gambling-related harms: a framework for action*. Birmingham, UK: Gambling Commission. doi: 0.1016/S2468-2667(24)00167-1

World Health Organization. (1996). WHOQOL-BREF: Introduction, Administration, Scoring, and Generic Version of the Assessment. Geneva, Switzerland: World Health Organization.

Article Submission: https://gamb.manuscriptmanager.net/