Abstract

While the PGSI is indeed an established index of problem-gambling symptoms, it nevertheless does not quantify the degree of harm experienced by individuals at different points on the spectrum of gambling problems. The purpose of the present study was to establish the relationship between the PGSI category and health-related quality of life (HRQoL) decrements using a population health (PH) method. Harms reported by gamblers and affected others across the PGSI spectrums were transformed into 798 vignettes. A general population panel (N=786) and experts who work with gamblers (N=51) rated the impact of these vignette descriptions on quality of life using the Time Trade-Off task, and a Visual Analogue Scale incorporating 27 comparison conditions. Disability weights (DW) were then estimated for different levels of gambling symptoms. A DW of 0.44 was estimated for problem gamblers (PG), suggesting a reduction in the effective enjoyment of life by over 4 years for every 10 years in lifespan. Lower—but non-negligible—DWs of .14 and .29 were determined for low- and moderate-risk gamblers. Gambling is compared with a number of other conditions with respect to HRQoL impact. On average, PG harm appears to be similar to that of a manic episode of bipolar disorder and severe alcohol abuse disorder. We discuss advantages, and methodological challenges, in applying PH methods to measuring the severity of gambling problems in terms of HRQoL.

Keywords: gambling harm, health related quality of life, disability weights, visual analogue scale, time trade-off, burden of disease
Résumé
Bien que l’indice du jeu excessif (PGSI) soit en effet un indice établi des symptômes liés aux problèmes de jeu, il ne quantifie pas le niveau de préjudice subi par les personnes situées à différents points sur le spectre des problèmes de jeu. Le but de l’étude a été d’établir la relation entre la catégorie PGSI et les écarts à la baisse en lien avec la qualité de vie liée à la santé (QVLS) en utilisant une méthode de santé de la population. Les torts signalés par les joueurs et les personnes touchées dans le spectre PGSI ont été transformés en 798 vignettes. Un groupe de population en général (N = 786) et des experts qui travaillent avec des joueurs compulsifs (N = 51) ont évalué l’incidence de ces descriptions de vignette sur la qualité de vie à l’aide de la tâche Time Trade-Off (marchandage de temps) et une échelle visuelle analogue intégrant 27 conditions de comparaison. Les poids d’incapacité (DW) ont ensuite été estimés pour différents niveaux de symptômes du jeu. Un DW de 0,44 a été estimé pour les joueurs compulsifs, ce qui laisse supposer une diminution de la jouissance réelle de la vie de plus de 4 ans pour chaque tranche de vie de 10 ans. Les DW inférieurs, mais non négligeables, de 0,14 et 0,29 ont été déterminés pour les joueurs à risque faible et modéré. Le jeu est comparé à un certain nombre d’autres conditions en ce qui concerne l’incidence de la qualité de vie liée à la santé (QVLS). En moyenne, le préjudice causé par un joueur compulsif s’apparente à celui d’un épisode maniaque de trouble bipolaire et d’un trouble sévère d’abus d’alcool. Nous discutons des avantages et des défis méthodologiques, en appliquant des méthodes de santé de la population pour mesurer la gravité des problèmes de jeu en termes de QVLS.

Introduction
Gambling has been recognized as an important public health concern that has a significant impact on population health (Abbott, Bellringer, Garrett, & Mundy-McPherson, 2014; Productivity Commission 2010; Wardle et al., 2011). Much research and policy development is devoted to assessing the negative impact of gambling on the community, and to reducing the prevalence and severity of gambling-related harm. Assessment of gambling harm, at the individual, group, and population levels, is critical for targeting resources to address gambling problems. While a number of studies have investigated the broad symptomology of problem gambling (Clarke et al., 2006; Davidson & Rodgers, 2011; Hounslow, Smith, Battersby, & Morefield, 2011; Li, Browne, Rawat, Langham, & Rockloff, 2016; Rodda, Lubman, & Latage, 2012; Subramaniam et al., 2015), a population measure of gambling harm is nonetheless currently lacking. Thus, a critical need exists for new measurement tools, not just to describe, but also to quantify the amount of gambling harm that is accruing to individuals and populations.
Since Korn and Shaffer (1999) first recommended adopting a public health approach to gambling nearly twenty years ago, a number of jurisdictions where gambling is legal have conducted regular prevalence surveys to determine patterns of gambling behaviour, gambler profiles, and to identify risk factors and comorbidities of interest (Abbott et al., 2014; Korn & Shaffer, 1999; Volberg et al., 2015; Wardle et al., 2011). However in conducting these surveys, the prevalence captured in the population of interest is a behavioural measure of gambling utilising a validated behavioural screening instrument such as the Problem Gambling Severity Index (PGSI) (Ferris & Wynne, 2001). The PGSI is a nine-item scale, ranging from 0–27, which measures the severity of gambling problems. The four types include non-problem gamblers (0), low-risk gamblers (1–2), moderate-risk gamblers (3–7), and problem gamblers (8–27). While such measures are indeed useful for service planning, they do not also allow population level health impacts of gambling to be compared to other public health priorities. To date no summary measures of gambling impact on population health have been developed which could support policy development and the prioritisation of health resources.

**A public health approach**

A public health approach encompasses population health, which seeks to understand why different groups within the population experience different health outcomes. The application of a public health approach, adopted by researchers and policy analysts internationally, has led to public health responses yielding significant reductions in morbidity and mortality, e.g., tobacco control, immunisation, road safety, and environmental contaminants control (Productivity Commission, 2010). One of the most widely used approaches for estimating the aggregate population-level impact of mortality and morbidity due to health conditions is the Global Burden of Disease (GBD) method, initiated in 1990 by the World Health Organization and the World Bank (Murray & Lopez, 1996). Over the last 25 years, burden of disease studies have estimated the impact on population health for hundreds of health states, ranging from diseases, cancers, substance use disorders, mental health disorders, and injuries (e.g., Mathers, Lopez, & Murray, 2006; Murray & Lopez, 1996; Salomon et al., 2012; Salomon et al., 2015; Stouthard, et al., 1997). Several common summary measures exist to calculate the burden of a health state, with the common factor involving utilisation of Health Related Quality of Life (HRQL) weights, known as disability weights (DWs), combined with the prevalence, incidence, severity, and/or duration of the health state. DWs are measured on a ratio scale between 0 and 1, with either 0 or 1 being equivalent to death, and the other end of the scale reflecting ideal health and well-being. The DW for a health state typically measures the decrement to quality of life it has on an individual living one year with that condition, and allows comparisons to be made between health states on their relative impact to the population within a given timeframe. For example, the Australian Burden of Disease Study (Australian Institute of Health and Welfare, 2015a) sourcing disability weights from the Global Burden of Disease Study 2010 (Salomon et al., 2012) and health and population data, calculated the burden of disease for 17 fatal and non-fatal disease groups in Australian
for the reference period of 2011. The study found that most of the burden of disease in 2011 was from chronic diseases, leading with cancer (19%), then cardiovascular diseases (15%), mental and substance use disorders (12%), and musculoskeletal conditions (12%) (AIHW, 2015b). These results are considered an important resource for health policy formulation, service planning and to monitor population health. The ability to calculate a disability weight for gambling related harm, and by level of gambling problems, would allow researchers and policy makers to estimate its burden to the Australian population and directly compare it to other prominent health conditions.

**Challenges to measuring the impact of gambling on population health**

Two significant challenges need to be overcome in developing a summary measure for gambling related health impacts. The first is the choice to be made in defining a case for gambling, and the second is the lack of robust measures needed to develop an appropriate disease model.

Defining a case for gambling is problematic because gambling is in fact a behaviour rather than a health outcome. However, pathological gambling is defined with diagnostic criteria by the International Classification of Diseases, 10th revision (ICD-10) (World Health Organization, 1992), and gambling disorder in the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) (American Psychiatric Association, 2013). While being diagnosed with pathological gambling or gambling disorder would logically count as a case, it is inadequate in terms of capturing the loss of population health from gambling because of two ways that it underestimates the impact. To meet the diagnostic standard for pathological gambling or gambling disorder, and therefore meet the requirement for being considered as a case, a person would already have experienced significant losses to health and wellbeing over a potentially extended period of time, as captured in the diagnostic description of persistent and recurrent (American Psychiatric Association, 2013). Furthermore the number of cases reported would be a significant undercount, based on very low treatment-seeking rates for gambling problems, and the use of non-clinical treatment and management options (Ledgerwood et al., 2013; Rigbye & Griffiths, 2011).

The alternate option is to focus on the behaviour, through what Hodgins, Stea, and Grant (2011) identified as problem gambling rather than pathological gambling. Problem gambling, while less formally defined, has been measured in numerous national and sub-national surveys in jurisdictions where gambling is legal (Abbott et al., 2014; Dowling et al., 2016; Wardle et al., 2011). The two most commonly used screens for problem gambling are the PGSI and the South Oaks Gambling Screen (SOGS), which are used in research, prevalence screens, and as an initial screen in some clinical contexts. The SOGS is used more by clinicians, and is mandated for use in some jurisdictions to capture data on levels of problem gambling (Holtgraves, 2009; Young & Wohl, 2009). The PGSI is designed for general rather than clinical use (Holtgraves, 2009), and tends to be used more in both population surveys and research contexts to capture gambling behaviour (Abbott et al., 2014; Hare, 2015;
Volberg et al., 2015; Wardle et al., 2011). Within the literature on gambling, in the absence of an alternate measure of harm, a conflation of gambling behaviour with gambling harm has taken place, with prevalence of PGSI categories reported as gambling harm (Langham et al., 2016). While the PGSI has established good psychometric characteristics (Orford, Wardle, Griffiths, Sproston, & Erens, 2010), its interpretation is categorical; i.e., discriminating between low-risk, moderate-risk, and problem gambling categories. Thus, it is not possible to infer that an individual with a PGSI score of 4 is probably suffering “twice the harm” as an individual with score 2. In statistical terminology, this limitation of the PGSI is summarized by the fact that it is measured on an ordinal rather than a metric scale. Despite these limitations, it is reasonable to infer that an individual’s experience of harm tends to increase with greater severity of gambling problems—as indicated by the PGSI.

To date there are no reliable data that captures the influence of gambling on mortality at a population level. Gambling has been examined in a number of studies in relation to its influence on suicide ideation (Battersby, Tolchard, Scurrah, & Thomas, 2006; Coroners Court of Victoria, 2013; Nower, Gupta, Blaszczynski, & Deverensky, 2004; Seguin et al., 2010) and suicide attempt or completion (Blaszczynski & Farrell, 1998; MacCallum & Blaszczynski, 2003; Newman & Thompson, 2007; Penfold, Hatcher, Sullivan, & Collins, 2006a; Penfold, Hatcher, Sullivan, & Collins, 2006b; Petry & Kiluk, 2002; Wong, Chan, Conwell, Conner, & Yip, 2010a; Wong, Chan, Conwell, Conner, & Yip, 2010b; Zangeneh & Hason, 2006). However, no systematic examination has been able to capture the contribution of gambling to mortality by suicide at a population level. Furthermore the role of gambling in contributing to other risk factors that cause premature mortality has not been examined through counterfactual analysis, although patterns of co-occurring risk factors (Goodwin, Browne, Rockloff, & Donaldson, 2015; Lloyd et al., 2010) and comorbidities have been identified (Cowlishaw, Merkouris, Chapman, & Radermacher, 2014; Holdsworth, Nuske, & Breen, 2013; Markham, Young, & Doran, 2012). In the absence of meaningful mortality data, the ability to estimate years of life lost to premature mortality is unachievable at this time.

Calculation of the loss of health to morbidity associated with engagement with gambling is also impacted through the lack of suitable metrics. A traditional disease model utilizes measures of incidence, duration and prevalence, although prevalence can be calculated from incidence and duration. Very few longitudinal studies have been conducted that include measures of gambling behaviours (Billi, Stone, Marden, & Yeung, 2014; Williams et al., 2015) and these have focussed on people who are already engaging in gambling. As highlighted above in terms of pathological gambling or gambling disorder, any reported incidence rates are significantly influenced by under reporting because of treatment-seeking behaviours of people experiencing harm from gambling. In terms of problem gambling, a four year longitudinal study of a randomly selected adult population (n = 15,000 W 1), identified an incidence rate (.036%) based on the PGSI category of problem gambler for the state of Victoria, in Australia (Billi et al., 2014). While the Victorian study offers some initial findings on the fluidity of gambling problems, and the move
between the levels of problem gambling levels, the data around duration of issues are still insufficient to determine a robust disease model.

The limitation of data available to inform a summary measure of gambling-related harm presents a number of challenges as outlined above. What is possible to determine for gambling would be a health state valuation for the impact of gambling-related harm on individual’s health. To date a health state valuation for gambling has not been developed, nor has gambling been examined as a risk factor in terms of its contribution to health outcomes. However, the absence of any appropriate or meaningful measure of impact places its own restrictions in terms of research and policy development for gambling. Cognisant of the limitations, but needing to progress the development of appropriate measures, it was determined that it would be possible to develop disability weights (DWs) that could quantify the loss of quality of life per year lived anchored to the PGSI categories. This links a standard health outcome measure (health state valuation) to a health behaviour (problem gambling status) captured in multiple population prevalence surveys. This allows for the calculation of annual morbidity component of the health gap created by gambling to be calculated for gambling.

**Aims**

The current paper deals with this objective of determining how much harm increasing gambling problems cause. This will be accomplished by conducting a health state valuation study that will relate scores on the PGSI to DWs. Our goal is to infer how much a typical individual is harmed, given a certain degree of gambling problems, in terms of the rate at which they accrue negative utility. The desired outcome is a table that relates raw PGSI scores to a DWs, a process which amounts to anchoring the PGSI to a bounded, metric scale measuring harm. Finally, the DWs corresponding to differing levels of gambling problems will be compared to DWs established for other major comparable health conditions.

**Method**

We shall describe an empirical approach for arriving at gambling DWs with respect to the PGSI, using preference based measures consistent with protocols recommended in Global Burden of Disease studies (Haagsma et al., 2015; Mathers, Vos, Lopez, Salomon, & Ezzati, 2001; Salomon et al., 2012) specifically using direct elicitation methods (Time trade off [TTO] and Visual analogue scale [VAS]), and using condition descriptions formed from an existing dataset of harm symptomology (Li et al., 2016). The method comprises three main components:

1. Formation of a corpus of vignettes describing the harm symptomology reported by individuals suffering from varying degrees of gambling problems.
2. Implementation of an online DW elicitation procedure, by which each vignette is evaluated by expert and general population raters using TTO and VAS protocols.
3. Analysis of the combined dataset, relating the elicited DW to the PGSI score of the individual reporting the harm symptomology.

**Vignettes describing harm symptomology**

Direct elicitation protocols for arriving at DWs require the development of stimuli describing the condition in question, for subsequent evaluation by participants. In other burden of disease DW studies, this process is often done via the researchers writing a small set of conditions descriptions, or vignettes, with content determined by clinical experts or research consensus (Bennett, Torrance, Boyle, Guscott, & Moran, 2000) for the sequela of interest. The method is more effective when the impact of the condition is relatively homogenous in the population, and the symptomology is relatively simple and well understood. In the case of gambling problems, this is arguably not the case, with symptomology being complex and variable, depending on the life circumstances and personal characteristics of the individual in question. Therefore, we adopted a formal procedure for creating a large set of vignettes intended to capture heterogeneity in the population experiencing a given level of gambling problems.

Recent research (Langham et al., 2016) reports on a systematic process for identifying and organizing the broad range of specific harms associated with gambling problems. From this work, a 73-item checklist was derived, comprising a comprehensive set of those harms, and organized within six broad domains (Li et al., 2016). Li and colleagues administered this checklist, and the PGSI, to 3076 gamblers and 2129 affected others. Over half of the gamblers reported gambling problems in the most severe category of the PGSI. Overall, this sample was primarily Australian born (80.5%), non-indigenous (97.1%), and recruited an approximately similar number of males and females. Details of the specific harms, their prevalence, and psychometric properties, can be found in the original article. To ensure our vignette stimuli represented adequately population heterogeneity in the experience of harm from gambling, we took a random sample of 798 cases, stratified by PGSI category, from this dataset (Table 1).

For each case, the specific harms nominated from the checklist were transformed to vignettes using a custom algorithm written for this purpose, generating natural language descriptors that closely mimicked the checklist text describing the original

<table>
<thead>
<tr>
<th>Vignette group</th>
<th>Problem Gambling Status</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Non-Problem</td>
<td>Low Risk</td>
</tr>
<tr>
<td>Own gambling</td>
<td>47</td>
<td>104</td>
</tr>
<tr>
<td>Affected others</td>
<td>18</td>
<td>29</td>
</tr>
<tr>
<td>Total</td>
<td>65</td>
<td>133</td>
</tr>
</tbody>
</table>
harms. The list of item-derived phrases, their allocation to sentences, and the sentence stems are provide in supplement 1. The vignettes were written in the first person, to suit the rating protocols (described in the following section), and in the case of affected others (i.e., persons harmed by another person’s gambling), nominated the relationship of the person to the gambler (e.g., father, spouse, sibling, etc.). The full set of vignette stimuli is provided in supplement 2.

**Online DW elicitation**

We utilized a direct framework to elicit the impact of gambling on one’s quality of life, rather than using an indirect approach and relying on existing instruments (e.g., the EurQol [EQ-5D]) (Hurst et al., 1994). Indirect approaches have some advantages in maximising comparability between conditions, but arguably lose sensitivity and validity when compared to direct evaluation of condition-specific descriptions (Rowen, Brazier, Tsuchiya, Young, & Ibbotson, 2012). This is particularly true in the case of gambling. For example, relationship dysfunction appears to play a major role in the experience of gambling harm (Li et al., 2016). However, relationships are not represented on the EQ-5D, potentially leading to an underestimate of harm when applied to gambling. Therefore, we adopted a directed method of utility estimation, by presenting participants with specific vignettes for evaluation.

We employed two protocols: an enhanced visual analogue scale (VAS), and a time trade-off (TTO) exercise. The VAS is a rating scale in which health states are evaluated by asking the respondent to indicate where on the scale from 0 (least harmful) to 100 (most harmful) he or she would place the health state. The DW is simply the rating from the VAS transformed into a value from 0 to 1. We modified the standard VAS to provide reference conditions (Salomon et al., 2012) on the visual scale, to encourage participants to evaluate the vignette with respect to conditions with established DWs. A screenshot of the interactive VAS tool is provided in supplement 3. Three unique VAS scales were developed, each with 9 reference conditions, selected to ensure a spread of health states that varied according to severity. A mouse-hover pop-up provided a more detailed description of the health condition. This information is summarized in supplement 4.

The TTO is another established protocol for eliciting DWs, in which duration of time spent living with impact from the condition is valued with respect to time spent free of the condition. The TTO provides a useful counterpoint to the VAS in that, while it is significantly more cognitively demanding to complete, it is more directly linked to the underlying health economic concept of utility (Whitehead & Ali, 2010). With reference to a 10-year period spent living with the condition, participants indicated how much lifetime they would give up in years, months and days, to avoid the harms described in the vignette. Supplement 5 shows a screenshot of the TTO task.

We recruited 786 participants (48% male) to evaluate the vignettes, with approximately half (391) evaluating a random sample of vignettes describing harms arising from one’s own gambling, and the remainder evaluating randomly selected vignettes...
Table 2

Breakdown of participants in the online elicitation protocols

<table>
<thead>
<tr>
<th>Participant Type</th>
<th>Vignette Group</th>
<th>Own Gambling</th>
<th>Affected Other</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gamblers</td>
<td></td>
<td>124</td>
<td>128</td>
<td>252</td>
</tr>
<tr>
<td>Affected Others</td>
<td></td>
<td>115</td>
<td>123</td>
<td>238</td>
</tr>
<tr>
<td>General population</td>
<td></td>
<td>123</td>
<td>122</td>
<td>245</td>
</tr>
<tr>
<td>Expert</td>
<td></td>
<td>29</td>
<td>22</td>
<td>51</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>391</td>
<td>395</td>
<td>786</td>
</tr>
</tbody>
</table>

from the point of view of an affected other. 20% of participants were 18–34 years old; 40% were aged between 35–54, and 40% aged 55+. A breakdown of participants is provided in Table 2. Individuals from four different populations participated in the current study. The first group included 51 “experts,” being gambling counsellors and support workers contacted through government funded organizations in Victoria. Each participant was offered a $40 gift card as compensation for his or her time. The remainder comprised approximately equal groups of gamblers, affected others, and the general Victorian population. The inclusion criteria for gamblers included having experienced harm from their own gambling. Affected others needed to have experienced harm because of someone else’s gambling, and inclusion for the general population group required them not to have experienced harm caused by their own or someone else’s gambling. These three groups were recruited and compensated by a commercial panel provider. We proceed by reporting two groups of interest: “general population” which combines the latter three groups, and “experts.” Overall, half the sample (50.4%) had contact in their personal lives with someone experiencing gambling-related harms, and this was higher for the expert sample (60.8%) than the general population (49.7%). Harm from their own gambling was reported by 17.6% of the sample, compared to 0 within the expert group. Lastly, experience harm from another’s gambling was experienced by 23.9% of the sample, and this was similar among both groups.

Each participant was provided with a link via e-mail to an online-hosted website with instructions, a consent form, and the interactive protocols. Each participant evaluated six vignettes via the VAS protocol, and then the same vignettes via TTO protocol. A total of 9,432 (612 expert) evaluations were elicited, with each of the vignettes being evaluated an average of 8.7 times (SD = 3.76).

Analyses

All analyses were conducted using the R statistical programming environment (R Core Team, 2008). VAS and TTO protocols both yield ratings that are interpretable as proportion, with (0,1) representing the maximal and minimal possible scores, combined and into a single dataset, and transformed to the logit scale prior in order to stabilize the error variance of the (0,1) bounded response. As TTO and VAS each composed half of the dataset, the overall analysis weighted each elicitation type
equally. We used isotonic simple regression (IR) (Barlow, Bartholomew, Bremner, & Brunk, 1972; Robertson, Wright, & Dykstra, 1988) to predict the elicited DW based on the PGSI associated with the vignette. IR provides for estimation of non-linear relationship between two variables, with the constraint that the relationship is monotonic and increasing. The simple regression provides a functional mapping between PGSI score, which is generally available from gambling prevalence surveys, and the average elicited DW. Standard errors were calculated by bootstrapping the estimated IR slopes with 1000 bootstrap replicates. Non-zero DW estimates were only calculated for vignette PGSI scores greater than zero. The goal of analysis was to establish a consensus evaluation of the DW-PGSI relationship. However, it was first necessary to check and account for intra-rater reliability, as well as “nuisance” participant and method effects.

We first checked for effects of participant characteristics on evaluations. We calculated means and standard errors with respect to PGSI category of the vignette, for expert and general population evaluations. As shown in Figure 1, while experts tended to provide lower DWs than others, these differences were not significant. The error bars within the figure indicate 95% confidence intervals of the mean.

Note that the larger error bars for the expert group are because of the relatively smaller sample size. We then regressed other participant characteristics on ratings. Females tended to provide slightly higher ratings ($B = .158, t = 2.98, p = .003$), as did those respondents testing positive on the lie-bet scale ($B = .196, t = 4.56, p < .001$). Affected others gave slightly lower ratings ($B = -.268, t = -3.03, p = .002$). However, these effects accounted for less than 4% of the variance in DW ratings. Method variance is an acknowledged issue in DW elicitation. As shown in Figure 2, VAS ratings tended to be higher than TTO ratings. However, a similar relationship between DW and PGSI was observed for both elicitation methods.

![Figure 1](image.png)

*Figure 1.* Disability weight for harm from one’s own gambling by rater group and PGSI category.
Health state valuation protocols are potentially a cognitively challenging task, requiring that participants have the capacity and the motivation to understand and follow the instructions. A potential disadvantage of Internet-based elicitation is that the researcher is unable to check personally for attention and understanding. Because the response variable is bounded between 0 and 1, if a reasonably large proportion of respondents are misunderstanding instructions, it has the potential to introduce bias into the estimates. Since we elicited DWs twice, for each condition description, for each participant, we could calculate a form of test-retest reliability for each participant in the form of a correlation between TTO and VAS ratings. 92% of the experts had a correlation of greater than 0.5. However, only 45% of the general population sample had correlations above 0.5. These included a small proportion of respondents with a strong negative correlation, suggesting that at least some participants had misunderstood the use of the TTO “slider,” and were responding in the opposite direction for this task relative to the VAS. Therefore, we considered schemes for down-weighting participants with less reliable estimates. According to Crocker and Algina (1986), “few, if any, standards exist for judging the minimum acceptable value for a test-retest reliability estimate” (p. 133), and determining what is acceptable requires a consideration of the cost of different types of measurement errors. It is also important to note that this is not a conventional use of the test-retest statistic, as (a) the retest involved both a different measure (VAS versus TTO) as well as a replication, (b) reliability was assessed within each rater’s six repeated scores, not for the dataset as a whole, and (c) our goal was not to validate an individual measure’s scores (i.e., an individual vignette rating), but rather only to confirm that minimal bias is affecting the sample averaged estimate from this source. We calculated candidate mean DW by PGSI using a logistic weighting function with a gradient of 5, and with correlation intercepts ranging from 0 to 0.6. The difference between the largest and smallest DW was <0.03 (on the [0,1] scale) for each PGSI category. This suggested that the mean estimates were robust to the test-retest reliability threshold, an arbitrary weighting

Figure 2. Disability weights for harm from one’s own gambling by elicitation method and PGSI score.
threshold of 0.3 was applied to ensure respondents satisfied a minimal criterion of reliability.

Within a given PGSI category (e.g., PG 8+), PGSI scores are not evenly distributed: e.g., lower scores tend to be more prevalent than higher scores. Thus, to arrive at DWs by category \( DW_{CAT} \), population weighting by relative PGSI score prevalence within categories is necessary to avoid bias. To arrive at a summary of DW with respect to PGSI category (as opposed to score), the DW by PGSI score values were population weighted with respect to the relative prevalence of each score within categories. This was done for each PGSI category \( CAT \) via

\[
DW_{CAT} = \sum_{PGSI=CAT_{MIN}}^{CAT_{MAX}} \frac{DW_{PGSI} \cdot N_{PGSI}}{N_{CAT}}
\]

using data from a recent Victorian prevalence survey (Hare, 2015).

**Results**

Figure 3 shows the estimated DW-PGSI relationship incorporating the full weighted dataset (error bars indicate 95% confidence intervals of the mean). Positive inflections in the DW are evident between gambling risk categories. This is notable, given that participants were not informed of the category of gambling problems associated with each vignette.

Table 3 summarizes the expected DW associated with each PGSI category.

Comparing harm to self-versus harm-to-others, gamblers in the low risk (DW = .13 versus .17) and moderate (.29 versus .33) categories appeared to “pass on” a similar (though slightly greater) quantity of harm to those around them. Those in the

![Figure 3. Disability weight for harm from one’s own gambling by PGSI score.](image-url)
problem gambler category caused significantly more harm to affected others (.36), but this was still significantly less than harm experienced personally (.44). Figure 4 compares DW for PGSI categories with other reference conditions, as sourced from the global burden of disease study (Salomon et al., 2012).

Problem gambling (.44) had a DW between moderate (.39) and severe alcohol use disorder (.55), and similar to bipolar disorder (manic episode) (.48). Moderate-risk gamblers (.29) had a DW similar to mild alcohol use disorder (.26) and stroke (moderate plus cognition problems) (.31). We estimated a DW for low-risk gambling problems (.13) as slightly less than a moderate anxiety disorder (.15).

Table 3

<table>
<thead>
<tr>
<th>PGSI Category</th>
<th>General Population</th>
<th>Experts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DW</td>
<td>Lower CI</td>
</tr>
<tr>
<td>Low Risk</td>
<td>0.14</td>
<td>0.12</td>
</tr>
<tr>
<td>Moderate Risk</td>
<td>0.29</td>
<td>0.27</td>
</tr>
<tr>
<td>Problem Gambler</td>
<td>0.46</td>
<td>0.43</td>
</tr>
</tbody>
</table>

Figure 4. Disability weights for harm from one’s own gambling compared to other health states.
Discussion

The deceptively simple question, “how much harm do gambling problems cause,” requires a detailed conceptual, empirical, and methodological base to be addressed effectively. Our approach builds on recent work to conceptualize (Langham et al., 2016) and comprehensively survey (Li et al., 2016) the broad range of specific harms that arise from gambling problems. We conducted a health state valuation study to integrate these diverse harms into a single metric, mapping PGSI scores to a proportional decrement to an individual’s quality of life.

The key contribution of the study relates to the DWs assigned to non-zero PGSI scores and (non-problem) categories. Essentially, the quality of life of problem gamblers is compromised/decreased by 44%, and by a smaller amount for moderate-risk gamblers (29%), and low-risk gamblers (13%). That is, most people would prefer to live only five and 1/2 years being free of gambling problems, rather than ten years with them. Our own judgement of gambling DWs with respect to other conditions (Figure 4) accord with these results: For example, the decrement of quality of life by severity of gambling problems appears aligned with the severity of alcohol use disorder, positioning gambling-related harm alongside that of alcohol related harm.

These figures validate and confirm the qualitative experiences reported by problem gamblers—that it represents a moderately severe level of psychological, social, and physical suffering. Though harm caused to low-risk gamblers is relatively mild by comparison, this figure is perhaps a more sobering finding. This finding validates the concern raised by the Productivity Commission (2010) that the aggregate of less severe harms experienced by a larger proportion of the population might be of more magnitude than the intensity of harm experienced by the small proportion of the population classified as problem gamblers. This is because low-risk gamblers are between 5–10 times more prevalent than problem gamblers—leading to the conclusion that the bulk of aggregate harm is actually accruing to individuals in lower-risk gambling categories gamblers.

Strengths and Limitations

Our approach had a number of strengths and some limitations. We used an online elicitation protocol, which allowed us to obtain a large and diverse sample, with differing perspectives on gambling harms. This was considered a robust method as it represented the views of those experiencing personal gambling problems, the community, and experts in the field. However, we found that many general population respondents did not provide reliable estimates. While this issue was overcome by adopting a statistical weighting criterion, it would have been better to implement a procedure to ensure participants were sufficiently motivated and trained to complete the protocols effectively. The biggest challenge is perhaps to ensure adequate and representative coverage of the population of individuals experiencing diverse levels and forms of gambling-related harm. Our approach incorporated a deliberately comprehensive set of harms, and algorithmic generation of a large number of
vignettes for evaluation, which were then stratified-sampled randomly within PGSI categories.

Scholars have criticized VAS for its evaluation of health-related quality of life. They have done so partly because of (1) perceived back of theoretical validity, on the part of VAS; and (2) recognition of the potential of scale biases, such as end-of-scale bias, where ratings at the extreme ends of the scale are avoided (Stubbs et al., 2000; Tolley, 2009; Whitehead & Ali, 2010). Mathers et al. (2001) recommend that a sufficient number of different reference states with different severity levels should be on the VAS to ensure respondents minimize scaling distortions. Our implementation of the VAS, which incorporates relative comparisons of the rater between the gambling vignette and comparison conditions, appears to address these concerns directly, and yields DW estimates for gambling problems that are positioned directly with other conditions.

The TTO, on the other hand, is consistent with economic models of decision-making and is fundamentally connected to the concept of utility (Whitehead & Ali, 2010). However, the TTO has been criticized for being too complex for many respondents (Dolan & Stalmeier, 2003; Rowen, Brazier, & Van Hout, 2014; Smith, Sherriff, Damschroder, Loewenstein, & Ubel, 2006). This may be why some researchers recommend that the TTO be administered face-to-face by an interviewer among small samples and among “experts,” such as health professionals (Norman et al., 2010; Shah, Lloyd, Oppe, & Devlin, 2013). Nevertheless, in many respects the VAS and TTO would appear to complement in each other in terms of their respective strengths and weaknesses.

**Future Work**

This study premised that non-problem gamblers (as assessed by the PGSI) could be assumed to be free of gambling-related harm. However, in light of the DWs determined with respect to low- and moderate-risk gamblers, there would appear to be a need to consider harms occurring to (purportedly) non-problem gamblers. A corollary to this, is the need for a dedicated short-form screen for assessing the occurrence of gambling harms in the population. The present study was restricted to harm accruing from current gambling behaviour, and has not assessed the DW for legacy harms (Langham et al., 2016) that continue to cause a decrement to health even if the behaviour ceases.

This study was conducted with an Australian sample. While it is probable that similar harms and DWs can be applied to similar countries, such as the US, Canada and New Zealand, this conjecture nevertheless still needs to be confirmed. This work focused on individual-level harm. Future work should apply burden of disease methods to integrate prevalence and severity information, to calculate the population-aggregate impact of gambling problems. Such projects will enable comparisons of gambling with similar social issues such as alcohol or depression, allowing for better-informed allocation of resources to harm minimisation.
Future research should focus on further efforts to refine and validate DWs for different levels of gambling problem severity. Acknowledged method variance falls between protocols, and intense research comparing and contrasting different protocols is ongoing (see e.g., Kularatna, Byrnes & Scuffham, 2014). Given a recent trend towards discrete choice protocols in DW estimation, researchers should consider applying these methods to harm from gambling.

A strength of the present study was employing a large set of vignette stimuli, intended to capture the fairly prevalence of harm symptomology across the spectrum of gambling problems. Nevertheless, it raises the question of whether it might be possible to capture a single “most typical” condition description associated with a given level of gambling problems. If this was accomplished, it would facilitate refinement and confirmation of the DWs that should be associated with this spectrum of conditions. Another promising alternative is to establish DWs not with respect to the PGSI, but rather with respect to a purpose-built population screen for gambling harm. This would enable direct measurement of harm severity, and facilitate the establishment of sensitive metric for monitoring aggregate harm in different populations.

Conclusion

This research generated a calculation of DWs for gambling problems (.44) that allows a direct comparison with other important population health concerns. The impact of problem gambling on a person’s quality of life is somewhere between that of moderate and severe alcohol abuse disorder. Moreover, the quantification of harm for problem gambling allows for a better future understanding of the population-level costs, and accordingly, the appropriate investment in harm minimisation. The method can provide a means for ongoing monitoring of harms because of gambling, and form the basis of a jurisdictional report card on progress towards harm reduction.

References


diseasemanual.pdf

mortality by condition: data, methods and results for 2001. In Lopez, Mathers,
Ezzati, M., Jamison, D. T., & Murray (Eds.). Global Burden of Disease and Risk

comprehensive assessment of mortality and disability from diseases, injuries and risk
factors in 1990 and projected to 2020. Cambridge, MA: Harvard School of Public
eng.pdf

gambling and attempted suicide: Findings from a national survey in Canada.

Does mode of administration matter? Comparison of online and face-to-face
administration of a time trade-off task. Quality of Life Research, 19, 499–508.
doi:10.1007/s11136-010-9609-5

Nower, L., Gupta, R., Blaszczynski, A., & Derevensky, J. (2004). Suicidality and
depression among youth gamblers: A preliminary examination of three studies.

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

attempted suicide: Part I: High prevalence amongst hospital admissions. International

attempted suicide: Part II: Alcohol abuse increases suicide risk. International Journal

Petry, N. M., & Kiluk, B. D. (2002). Suicidal ideation and suicide attempts in
treatment-seeking pathological gamblers. Journal of Nervous and Mental Disease,
190, 462–469. doi:10.1097/01.NMD.0000022447.27689.96


WHAT IS THE HARM?


Submitted August 14, 2016; accepted February 14, 2017. This article was peer reviewed. All URLs were available at the time of submission.

For correspondence: Matthew Browne, Ph.D., B8 G.47 University Drive, Branyan, QLD, 4670, Australia. E-mail: m.browne@cqu.edu.au

Competing interests: None declared (all authors).

Ethics approval: The Central Queensland University Human Research Ethics Committee approved this study on the 19th of June 2015 (H15/05-126).

Acknowledgements: This research was funded by a grant from the Victorian Responsible Gambling Foundation.

**Supplementary Material**

**Supplementary File 1.** Item-derived phrases, sentence allocations, and sentence stems

**Supplementary File 2.** Vignette Stimuli

**Supplementary File 3.** Visual analogue task

**Supplementary File 4.** Health state reference conditions

**Supplementary File 5.** Time trade-off task