INTERNATIONAL GAMBLING STUDIES, 2017
http://dx.doi.org/10.1080/14459795.2017.1284250

Routledge
Taylor & Francis Group

Use of online crowdsourcing platforms for gambling research

Sandeep Mishra[a] and R. Nicholas Carleton[b]

[a]Faculty of Business Administration, University of Regina, Regina, Canada; [b]Anxiety and Illness Behavior Laboratory, University of Regina, Regina, Canada

ABSTRACT

Crowdsourcing platforms like Amazon's Mechnical Turk and Crowdflower have been touted to be a cost-effective way to collect large amounts of behavioural data. Across four large-*n* studies, gambling-related behaviours, tendencies and traits among participants in these labour markets were examined. In Studies 1 and 2, both conducted on Crowdflower, problem gamblers (as measured by the benchmark Problem Gambling Severity Index) comprised 24.5% and 21.9% of participants, respectively. In Study 3, conducted on Mechanical Turk, problem gamblers comprised 9.0% of participants. In Study 4, a two-wave longitudinal study conducted on Crowdflower, problem gamblers comprised 13.5% of participants in wave one and 14.8% of participants in wave two. In Studies 2 and 3, strong convergent associations were demonstrated across various measures of problem gambling tendencies and general gambling involvement. Furthermore, it was demonstrated that gambling was associated with personality traits (impulsivity, sensation-seeking, self-control), risk attitudes, affect, and behavioural risk-taking consistent with previous research. In Study 4, it was demonstrated that measures of problem gambling have acceptable test-retest reliability. Online crowdsourcing platforms appear to offer access to samples with remarkably high proportions of problem gamblers. However, this characteristic means that such samples are not necessarily representative of gambling tendencies among more general populations.

ARTICLE HISTORY
Received 9 January 2016
Accepted 14 January 2017

KEYWORDS
Methodology; sampling; data collection; crowdsourcing; gambling; risk taking

Some of the most difficult aspects of behavioural research involve sampling, recruiting, and collecting data from representative participants. Many studies rely on convenience samples (e.g. undergraduate students) and even studies involving more diverse community members have shortcomings (e.g. sampling from a single local population limits generalizability). Online crowdsourcing platforms offer a potential solution to the difficulties of recruiting and compensating representative participants. Here, we examine whether crowdsourcing is specifically useful for gambling researchers. In the following, we (1) describe two of the most commonly used crowdsourcing markets – Amazon Mechanical Turk and Crowdflower, (2) review characteristics of crowdsourced samples, and (3) present four studies examining the suitability of crowdsourced samples for gambling research.

CONTACT Sandeep Mishra mishrs@gmail.com

What is crowdsourcing?

Crowdsourcing platforms are online labour markets. *Customers* or *requesters* post computerized tasks for *workers* or *contributors* to complete. Among the most popular markets are Amazon's Mechnical Turk (MTurk) and Crowdflower. MTurk is a single marketplace run by multinational company Amazon involving over half-a-million potential workers (at last estimate; Paolacci & Chandler, 2014). MTurk workers are compensated directly through cash payments. Crowdflower is an aggregating service that sources contributors from various other online labour channels. Crowdflower compensation depends on the channel from which people are recruited. Some channels compensate workers in cash; others compensate with such different currencies as raffle tickets, draws, or proprietary site-specific currencies. Crowdflower advertises access to over five million workers across all channels.

Tasks provided to workers or contributors on MTurk and Crowdflower are typically short and inexpensive (usually no longer than 30 minutes, and compensation is usually less than a dollar; Buhrmester, Kwang, & Gosling, 2011). Data can be collected from hundreds (or even thousands) of participants in a matter of days. Any task that can be computerized can be presented, including, but not limited to, surveys, behavioural tasks, and experimental tasks. Workers/contributors can browse among available tasks and self-select into tasks they find interesting, desirable or relevant.

For the purposes of behavioural research, crowdsourced tasks typically involve a link to an external platform for research (e.g. Qualtrics). Participants complete a task and receive a code, which they input into MTurk/Crowdflower to receive immediate compensation. Compensation can be refused if a task has been done poorly, which also has reputational consequences; each worker/contributor has a site-wide rating of the proportion of tasks that they have completed satisfactorily and customers/requesters can limit workers/contributors based on their reputation.

Characteristics of crowdsourced samples

Crowdsourced participants are motivated by both intrinsic and extrinsic factors. Although some people participate in studies for monetary reasons, others report doing so because the tasks themselves are enjoyed (e.g. Buhrmester et al., 2011; Paolacci & Chandler, 2014). The mix of both intrinsic and extrinsic motivations is important given that workers complete tasks for small amounts of pay, potentiating questions at baseline about data quality.

Crowdsourced participants are generally more demographically representative of the general population than typical convenience samples of undergraduate students (reviewed in Paolacci & Chandler, 2014). However, some systematic participant characteristics have been observed. Crowdsourced samples over-represent Internet users relative to the general population. Consequently, crowdsourced samples are typically younger, more educated, less religious, and more liberal than the general population (Paolocci & Chandler, 2014). Compared to undergraduate and community samples, people in crowdsourced samples are also more socially anxious and show greater levels of psychopathology (Shapiro, Chandler, & Mueller, 2013). Crowdsourced samples also tend to exhibit lower levels of extraversion and emotional stability (Goodman, Cryder, & Cheema, 2013). However, the aforementioned sample differences are relatively small and crowdsourced samples are more representative than most convenience samples used in behavioural research.

Previous research spanning several disciplines and domains has demonstrated the reliability and validity of behavioural data from crowdsourced participants. Paolocci, Chandler, and Ipeirotis (2010) demonstrated that classic experimental results in the judgment and decision-making literature (e.g. framing effects; heuristics and biases) replicate consistently on MTurk with effect sizes comparable to other samples. Horton, Rand, and Zeckhauser (2011) similarly evidenced MTurk as an excellent platform for experimental economics research. Crowdsourcing has also been recently used for gambling research (e.g. Hollingshead, Kim, Wohl, & Derevensky, 2016; Hoon & Dymond, 2013; Kim & Hodgins, in press; Kim, Wohl, Salmon, Gupta, & Derevensky, 2015; Tabri, Dupuis, Kim, & Wohl, 2015). More broadly, crowdsourcing has been used for research in political science (Berinsky, Huber, & Lenz, 2012), cognitive psychology (Crump, McDonnell, & Gureckis, 2013), linguistics (Fort, Adda, & Cohen, 2011; Sprouse, 2011), and clinical psychology (Beshai, Mishra, Meadows, Parmar, & Huang, 2017; Shapiro et al., 2013), among other areas (reviewed in Paolocci & Chandler, 2014).

Successful use of crowdsourcing platforms for clinical psychology research is particularly important to gambling researchers. Shapiro and colleagues (2013) demonstrated that widely used mental health measures showed satisfactory internal reliability and test-retest reliability in two MTurk samples. Furthermore, they demonstrated criterion validity of these screens by associating psychopathology with established predictors (e.g. unemployment). Finally, prevalence rates of depression, anxiety, and trauma exposure among MTurk samples largely matched prevalence rates in the general population. Such clinically relevant results underscore the potential of crowdsourcing samples for gambling researchers. Past-year problem gambling prevalence estimates in country-level populations range from 0.5% to 7.6%, with an average prevalence rate of 2.3% (Williams, Volberg, & Stevens, 2012). Crowdsourcing offers an opportunity to more easily collect data from such individuals.

Overview

We present four studies examining the suitability of crowdsourcing platforms for gambling research. Studies 1 and 2 were conducted on Crowdflower. Study 1 examined reported problem gambling tendencies (using the widely used Problem Gambling Severity Index; PGSI) and frequency of general gambling behaviours, as well as associations of gambling with demographic variables known to be correlates. Study 2 examined whether problem gambling tendencies were associated with a key individual differences measure (self-control), risk attitudes, and behavioural risk-taking (all previously associated with gambling; reviewed in Toneatto & Nguyen, 2007; van Brunschot, 2009). Study 3 extended Studies 1 and 2 by examining problem gambling tendencies and general gambling involvement on a different crowdsourcing platform (MTurk) by including an additional measure (the Problem and Pathological Gambling Measure; PPGM), and by examining associations of gambling with additional relevant individual differences (sensation-seeking, impulsivity, personal relative deprivation). Study 4 examined the test-retest reliability of two important problem gambling measures (PGSI, PPGM) using a two-wave longitudinal design on Crowdflower.

These studies extend previous work examining the suitability of crowdsourcing for behavioural research in several ways. First, we focus specifically on gambling-related outcomes and behaviour. Although some research has previously examined clinical psychopathological outcomes among crowdsourced samples (e.g. depression and depressive symptoms, anxiety, trauma exposure; Beshai et al., 2017; Shapiro et al., 2013), little research has specifically

examined problem gambling tendencies or more general gambling behaviours (but see Kim & Hodgins, in press). Second, we examine outcomes and behaviour on both Crowdflower and MTurk. To our knowledge, no previous studies have specifically addressed the suitability of Crowdflower as a recruiting platform. This shortcoming is particularly important given that at the time of writing, MTurk was no longer available to researchers outside the United States. Third, we examine associations between problem gambling tendencies and general gambling involvement alongside a wide array of personality traits, attitudes, affective outcomes, and behavioural outcomes.

Study 1

Method and measures

A total of 493 participants (262 female, 230 male, 1 trans*; age: $M = 37.1$, $SD = 12.9$, Range: 18 to 76) were recruited from Crowdflower. The title of the study displayed to potential participants was 'A Short Survey on Decision-Making'. Being a gambler was not an inclusion criterion for this or any of the other studies. The formal introduction and consent form to the study noted in general terms that participants would be completing measures of gambling (among others), but gambling measures were not explicitly highlighted. After completion of all measures, participants received a full debriefing including a detailed description of study aims. The three subsequent studies reported in this article used the same recruitment procedures.

Participants completed several demographic measures, including country, age, gender, employment status, and highest educational attainment. Participants were compensated with $0.50 USD or the equivalent (i.e. translated into site-specific currencies for some Crowdflower channels). Formal ethics approval for this study (and all others in the article) was granted by the University of Regina Research Ethics Board. All participants provided written informed consent and were thoroughly debriefed following participation in all studies.

General gambling involvement

General gambling involvement was measured through self-report of frequency of engagement in 12 different gambling activities in the last year: sports lotteries/pools; other lottery/instant win tickets; charity raffles/fundraisers; bingo; VLTs; slot machines; casino table games; horse race betting; poker for money in a public facility; poker for money at home; betting on non-poker games of skill; online gambling. Participants reported how frequently they engaged in each behaviour, on average, on the following scale: 1 (*never*), 2 (*less than 1 time per year*), 3 (*1–5 times per year*), 4 (*6–11 times per year*), 5 (*1 time per month*), 6 (*2–3 times per month*), 7 (*once a week or more*). The two dependent measures were (1) diversity of gambling activities engaged in and (2) mean frequency of engagement across all gambling activities.

Problem gambling tendencies

Problem gambling tendencies were measured using the Problem Gambling Severity Index (PGSI; Ferris & Wynne, 2001). The PGSI measures frequency of nine outcomes and behaviours associated with disordered gambling (e.g. 'Have you felt that you might have a problem with gambling'). Items were rated on a scale from 0 (*never*) to 3 (*almost always*); higher

scores indicated greater problem gambling tendencies. For categorical analyses, we used the following clinically informed PGSI cut-off scores recommended by Williams and Volberg (2014): non-gamblers or non-problem gamblers (0), at-risk gamblers (1–4) and problem and pathological gamblers (5–27). The PGSI is one of the most widely used measures of problem and pathological gambling. The measure is highly reliable with excellent external and internal validity (e.g. Orford, Wardle, Griffiths, Sproston, & Erens, 2010; Stinchfield et al., 2016), and is cross-culturally robust (e.g. Arthur et al., 2008; Colasante et al., 2013; Loo, Oei, & Raylu, 2011); however, the measure does not distinguish non-gamblers and non-problem-gamblers.

Results and discussion

Demographics

Participant recruitment was restricted to the Anglosphere – the set of western, English-speaking countries that have a similar cultural heritage (i.e. Australia, Canada, Ireland, New Zealand, the United Kingdom, the United States). Our Crowdflower sample had substantial variability in age and was older than typical undergraduate convenience samples. Descriptive statistics for employment status and highest level of educational attainment are provided in Table 1.

Gambling

Descriptive statistics for general gambling involvement were as follows: diversity of gambling activities ($M = 4.60$, $SD = 3.63$, *median* = 4.00); mean frequency of gambling activity ($M = 1.84$, $SD = 1.00$, *median* = 1.50), where 1 = *never*, 2 = *less than once per year*, 3 = *1 to 5 times per year*, 4 = *6 to 11 times per year*, 5 = *once per month*, 6 = *2 to 3 times per month*, 7 = *once or more per week*. The proportion of participants who reported not gambling at all in the last year was 12.2%. The reliability of the PGSI measure was high (*Cronbach's* $\alpha = .96$). Descriptive statistics for the PGSI were as follows: $M = 1.93$; $SD = 4.42$. Participants were categorized based on PGSI scores as follows: non-problem gamblers or non-gamblers (59.8%); at-risk gamblers (15.6%); problem gamblers (24.5%).

One of the biggest factors associated with gambling is being male (Raylu & Oei, 2002). In our sample, men scored significantly higher on the PGSI than did women, $t(431) = 2.71$, $p = .007$, $d = .26$; men: $M = 2.51$, $SD = 4.94$; women: $M = 1.37$, $SD = 3.74$. Men also reported engaging in more different types of gambling activities than did women, $t(489) = 4.38$, $p<.001$, $d = .39$; men: $M = 5.35$, $SD = 3.98$; women: $M = 3.94$, $SD = 3.13$. Finally, men reported significantly greater mean frequency of general gambling behaviour than did women, $t(488) = 4.18$, $p<.001$; men: $M = 2.03$, $SD = 1.18$, $d = .37$; women: $M = 1.66$, $SD = .76$.

Another key factor associated with gambling is age (Raylu & Oei, 2002); younger people typically engage in higher levels of gambling behaviour and exhibit higher levels of problem gambling tendencies. Older people showed significantly fewer problem gambling tendencies, $r = -.18$, $p<.001$. Neither diversity of gambling activities ($r = .01$, $p = .80$) nor frequency of general gambling involvement ($r = -.07$, $p = .12$) was associated with age.

The proportion of problem gamblers in the current sample was remarkably high. The observed rate in this Crowdflower sample (24.5%) is higher than that observed among prison populations (Williams, Royston, & Hagen, 2005) and among male undergraduates (e.g. Mishra, Morgan, Lalumière, & Williams, 2010); both of these populations show

Table 1. Descriptive statistics across Studies 1 to 4.

	Study 1	Study 2	Study 3	Study 4
Employment status				
Employed full-time	48.7%	45.1%	51.4%	54.5%
Employed part-time	14.1%	16.2%	21.0%	15.5%
Unemployed	31.7%	13.8%	23.8%	23.0%
Homemaker		11.2%		
Student		9.2%		
Retired	5.5%	3.5%	3.6%	7.0%
Highest level of education				
Did not finish high school	1.4%	2.4%		1.5%
High school or GED	22.0%	17.4%	8.4%	19.0%
Some college/university	28.5%	25.5%	35.6%	22.0%
College/university	36.0%	33.2%	43.3%	43.0%
Graduate/postgraduate/professional degree	12.0%	15.5%	12.8%	14.5%
Relationship status				
Single		35.2%		34.0%
Dating/long-term relationship		13.9%		7.0%
Married or equivalent		38.5%		51.0%
Separated/divorced		5.5%		5.5%
Widowed		1.4%		2.5%
Not reported		5.4%		
Ethnicity				
Aboriginal		2.6%		
African-American		3.6%		
Asian		8.4%		
Caucasian		78.8%		
Other/not reported		9%		
Annual income				
Mean			$37,847	
SD			$38,819	
<$10,000		6.5%		
$10,000 to $19,999		11.4%		
$20,000 to $29,999		12.4%		
$30,000 to $39,999		11.0%		
$40,000 to $49,999		8.5%		
$50,000 to $59,999		11.5%		
$60,000 to $69,999		7.1%		
$70,000 to $79,999		7.1%		
$80,000 to $89,999		5.0%		
$90,000 to $99,999		4.9%		
$100,000 to $149,999		7.20%		
$150,000+		1.80%		
(Not reported)		5.50%		

elevated high rates of gambling compared to the general population (reviewed in Williams et al., 2012). We also observed that problem gamblers comprised more of the sample than did at-risk gamblers, an unexpected result. These results may be due to the Crowdflower recruitment strategy. Crowdflower utilizes multiple channels and each involves different payout currencies. Some channels pay participants in cash, whereas others use raffle tickets or proprietary 'points' that can be used to purchase goods. People with problem gambling tendencies may be particularly drawn to online services that essentially offer legalized gambling opportunities as compensation.

Study 2

Study 2 was conducted using Crowdflower. In addition to measures of problem gambling tendencies, we included measures of self-control, risk attitudes and behavioural risk-taking.

These measures have all been associated with gambling in previous research (e.g. Mishra et al., 2010, 2011, in press). Including these measures allowed us to examine whether crowd-sourced samples demonstrated gambling-related associations comparable to convenience samples and/or community samples.

Method and measures

A total of 1386 participants (705 women, 596 men, 2 trans*, 83 unreported gender) were recruited from Crowdflower. Participants completed several demographic measures: gender, relationship status, household income, highest educational attainment, employment status, and race. Age did not record properly due to a coding error. As in Study 1, participation was restricted to the Anglosphere. Participants were compensated with $0.75 USD or the equivalent.

Problem gambling tendencies

Problem gambling tendencies were measured using the PGSI (as in Study 1). A general gambling involvement measure was not included in this study.

Self-control

Trait self-control was measured using the *Brief Self-Control Scale* (BSCS). The BSCS (Tangney, Baumeister, & Boone, 2004) consists of 13 items measuring self-control (e.g. 'I often act without thinking through all the alternatives') rated on a scale from 1 (*not at all like me*) to 5 (*very much like me*).

Subjective relative deprivation

Feelings of relative deprivation – an affective measure of the degree to which people feel subjectively deprived relative to others – were measured using the *Personal Relative Deprivation Scale* (PRDS). The PRDS consists of four items (e.g. 'When I think about what I have compared to others, I feel deprived'; Callan, Ellard, Will Shead, & Hodgins, 2008) rated on a scale from 1 (*strongly disagree*) to 7 (*strongly agree*). This measure has previously been associated with problem and pathological gambling tendencies (Callan, Will Shead, & Olson, 2015; Callan et al., 2008; Mishra & Novakowski, 2016).

Risk attitudes

Risk attitudes in six content domains (i.e. ethical, financial, health/safety, social, gambling and recreational) were measured using the revised *Domain-Specific Risk-Attitudes Scale* (DOSPERT; Blais & Weber, 2006). Likelihood of engagement in various risk behaviours was rated on a scale from 1 (*extremely unlikely*) to 7 (*extremely likely*).

Behavioural risk-taking

Behavioural risk-taking was measured using the *Choice Task* (CT; adapted from Fessler, Pillsworth, & Flamson, 2004; Mishra & Lalumière, 2010). Participants made six choices between pairs of monetary options. Both options had equal expected values but differed in outcome variance (e.g. 'Would you rather choose [A] $3 guaranteed, or [B] a 10% chance of earning $30?'). A total score of number of risky choices was computed.

Results and discussion

Demographics

Descriptive statistics for employment status, highest level of education, relationship status, ethnicity, and annual income are provided in Table 1.

Gambling

The reliability of the PGSI measure was high (*Cronbach's* $\alpha = .94$). Descriptive statistics for the PGSI were as follows: $M = 1.05$; $SD = 2.76$. Participants were categorized as follows: non-problem gamblers or non-gamblers (68.4%); at-risk gamblers (9.7%); problem gamblers (21.9%). The proportion of problem gamblers in the current sample was also remarkably high. Again, Crowdflower participants compensated through raffle tickets or other gambling currency analogues may be disproportionately likely to demonstrate gambling problems due to a self-selection bias.

Gambling associations

Men scored significantly higher on the PGSI than did women, $t(1211) = 7.55$, $p<.001$, $d = .43$; men: $M = 1.70$, $SD = 3.25$; women: $M = 0.52$, $SD = 2.14$. Correlations between PGSI scores and self-control (BSCS), behavioural risk-taking (CT), relative deprivation (PRDS), overall risk attitudes (total DOSPERT score), and self-rated physical health are provided in Table 2. Correlations between PGSI scores and domain-specific risk attitudes are presented in Table 3. Problem gambling tendencies were associated with all individual differences as expected, consistent with previous findings. Specifically, PGSI scores were associated with lower trait self-control, greater behavioural risk-taking, greater feelings of personal relative deprivation, and greater risk-prone attitudes. Together, the results of Study 2 provide further evidence suggesting that Crowdflower samples include participants who exhibit a wide range of problem gambling tendencies.

Study 3

Study 3 extended Studies 1 and 2 in three ways. First, we examined problem gambling tendencies and general gambling involvement among participants on a different crowdsourcing platform, Amazon Mechanical Turk (MTurk). Second, we incorporated an additional measure of problem gambling behaviour (the PPGM; Williams & Volberg, 2010), and measures of trait sensation-seeking and impulsivity. Third, we included an attention check item specifically designed for online crowdsourcing participants (Goodman et al., 2013).

Table 2. Correlations between PGSI scores and relevant individual differences.

	BSCS	CT	PRDS	DOS-T
PGSI	−.23***	.17***	.12***	.62***
BSCS		−.07*	−.34***	−.32***
CT			.05	.22***
PRDS				.17***

Notes: *$p<.05$; ***$p<.001$; PGSI = Problem Gambling Severity Index score; BSCS = self-control; CT = choice task; PRDS = relative deprivation; DOS-T = risk attitudes total score.

Table 3. Correlations between PGSI scores and domain-specific risk attitudes.

	DOS-S	DOS-R	DOS-G	DOS-I	DOS-HS	DOS-E
PGSI	.15***	.49***	.63***	.38***	.54***	.64***
DOS-S		.30***	.15***	.31***	.37***	.24***
DOS-R			.60***	.52***	.65***	.62***
DOS-G				.52***	.66***	.75***
DOS-I					.43***	.47***
DOS-HS						.76***

Notes: ***$p<.001$; PGSI = Problem Gambling Severity Index score; DOS-S = social risk attitudes; DOS-R = recreational risk attitudes; DOS-G = gambling risk attitudes; DOS-I = investment risk attitudes; DOS-HS = health/safety risk attitudes; DOS-E = ethical risk attitudes.

Method and measures

A total of 815 participants (403 women, 348 men, 3 trans*, 61 unreported gender; age: $M = 34.2$, $SD = 12.3$, Range: 18 to 76) were recruited from MTurk. Participants were compensated with $0.75 USD. As in Studies 1 and 2, participation was restricted to the Anglosphere. Participants completed (a) several demographic measures: age, gender, relationship status, an estimate of household income, highest educational attainment, and employment status; (b) the following individual differences and behavioural measures administered in Studies 1 and/or 2: PGSI, BSCS, CT, PRDS, DOSPERT; (c) additional measures as indicated below. Participants also completed an attention check as described in Goodman et al. (2013).

Impulsivity

Trait impulsivity was measured using *Eysenck's Impulsivity Scale* (EIS; Eysenck, Pearson, Easting, & Allsopp, 1985), which consists of 19 yes/no statements about impulsive behaviours (e.g. 'Do you often buy things on impulse?'). A total score was obtained by summing the number of 'yes' answers; higher scores indicated greater impulsivity.

Sensation-seeking

Sensation-seeking was measured using the *Brief Sensation-Seeking Scale* (BSSS; Hoyle, Stephenson, Palmgreen, Lorch, & Donohew, 2002), which consists of eight items measuring preferences for varied, stimulating experiences and disinhibited behaviour (e.g. 'I would love to have new and exciting experiences, even if they are illegal'). Items were assessed on a scale from 1 (*strongly disagree*) to 5 (*strongly agree*). Items were summed with higher scores indicating greater sensation-seeking.

General gambling involvement

General gambling tendencies were assessed using the *Gambling Behavior Scale* (GBS; Mishra, Lalumiere, & Williams, 2010), which assesses self-reports of frequency of engagement in various gambling behaviours, as well as the diversity of different gambling behaviours engaged in (modified to measure engagement both in the last year, and in the last month). Mean frequency of gambling for each behaviour was coded categorically as follows: 1 = *never*; 2 = *1–11 times total*; 3 = *1 time per month*; 4 = *2–3 times per month*; 5 = *1 time per week*; 6 = *2–6 times per week*; 7 = *daily*.

Problem gambling tendencies

Problem gambling tendencies were again measured with the PGSI. We also measured problem gambling tendencies using the Problem and Pathological Gambling Measure (PPGM; Williams & Volberg, 2010). The PPGM is a relatively new measure that has been demonstrated to have excellent reliability, validity, and problem gambling classification accuracy (e.g. Williams & Volberg, 2014; Williams et al., 2012). The PPGM consists of 14 'yes/no' items and provides 3 subscale scores; specifically, *gambling problems* (i.e. harm to self or others from gambling), *impaired control* (e.g. chasing losses), and *other issues* associated with gambling (e.g. rumination on gambling).

Results and discussion

Data preparation

Of the 815 participants, 60 exited the survey before reaching the debriefing page; data from these participants were not included in the results presented below. Among the 755 remaining participants, 700 (92.7%) correctly completed the attention check item. Participants who failed the attention check were not included in the results reported below.

Demographics

Descriptive statistics for employment status, highest level of education, and annual income are provided in Table 1.

Gambling

Descriptive statistics for general gambling involvement were as follows: diversity of gambling activities in the last month (M = 1.94, SD = 2.61, *median* = 1.00); diversity of gambling activities in the last year (M = 3.21, SD = 3.05, *median* = 2.00); frequency of gambling activity in the last month (M = 6.19, SD = 10.57, *median* = 0); frequency of gambling activity in the last year (M = 13.17, SD = 13.64, *median* = 10.00).

The reliability of the PGSI measure was high (*Cronbach's* α = .95). Descriptive statistics for the PGSI were as follows: M = 1.35; SD = 3.80. Participants were categorized as follows: non-problem gamblers or non-gamblers (75.7%); at-risk gamblers (15.3%); problem and pathological gamblers (9.0%). The proportion of problem gamblers in this sample was also remarkably high compared to population prevalence estimates, although lower than in Studies 1 and 2.

The reliability of the PPGM measure was high (*Cronbach's* α = .92). Descriptive statistics for the PPGM were as follows: total score: M = .55, SD = 1.70; problems subscale: M = .18; SD = .76; impaired control subscale: M = .22; SD = .63; other issues subscale: M = .14, SD = .53. Participants were categorized as follows: non-gamblers (84.4%); recreational gamblers (4.7%); at-risk gamblers (5.4%); problem gamblers (1.4%); pathological gamblers (3.9%).

Gambling associations

Men scored significantly higher on the PGSI than did women, $t(695) = 6.00$, $p<.001$, $d = .45$; men: M = 2.25, SD = 4.80; women: M = 0.57, SD = 2.39. Men also scored significantly higher on the PPGM (overall score) than did women, $t(694) = 5.68$, $p<.001$, $d = .43$; men: M = .93, SD = 2.20; women: M = .21, SD = .98. Men scored higher than women on all three subscales of the PPGM as well (all ts>4.84, ps<.001, ds>.37). All four measures of general gambling

involvement also showed significant sex differences, with men scoring higher than women on all measures (all *t*s>5.32, *p*s<.001, *d*s>.40).

Older people exhibited fewer problem gambling tendencies as measured by the PGSI (r = -.085, p = .025) and the PPGM (total score; r = -.078, p = .040). Age was associated with the PPGM problems subscale (r = -.092, p = .015), but not the impaired control or other issues subscales (both *r*s>-.06, *p*s>.10). None of the general gambling involvement measures (diversity of activities or frequency in the last year/month) were associated with age (all *r*s>-.059, *p*s>.12).

Correlations between all the gambling measures are provided in Table 4. All gambling measures were significantly associated. Correlations between the gambling measures (PGSI, PPGM, GBS), and trait self-control (BSCS), trait impulsivity (EIS), trait sensation-seeking (BSSS), behavioural risk-taking (CT), relative deprivation (PRDS), and overall risk attitudes (total DOSPERT score) are provided in Table 5. Correlations between PGSI scores and domain-specific risk attitudes are presented in Table 6. Problem gambling tendencies were correlated with all individual differences as expected, consistent with previous results. Diversity of gambling activities engaged in was not associated with the CT or the PRDS, although general gambling involvement frequency in the last year did correlate with both the CT and the PRDS.

Study 4

Study 4 examined three-to-four-week test-retest reliability of the PGSI and the PPGM on Crowdflower. Previous studies on MTurk have shown one-week test-retest values for mental health measures to range from .68 to .97 (Shapiro et al., 2013). We expected test-retest values for the PGSI and PPGM to be similar.

Method and measures

Data were collected in two waves from Crowdflower. Participation was again restricted to the Anglosphere. The first wave was open for one week of data collection; the second wave was administered three weeks later, and was also open for one week of data collection. Consequently, we assessed three-to-four-week test-retest reliability. Participants received $0.20 USD (or equivalent) for completing the first wave, and $0.60 USD (or equivalent) for completing the second wave. A total of 452 participants (244 women, 205 men, 3 trans*,

Table 4. Correlations between various measures of gambling and problem gambling tendencies.

	PPGM-T	PPGM-P	PPGM-I	PPGM-O	GBS-AY	GBS-AM	GBS-FY	GBS-FM
PGSI	.79***	.74***	.66***	.70***	.39***	.47***	.35***	.49***
PPGM-T		.90***	.87***	.87***	.42***	.50***	.35***	.49***
PPGM-P			.63***	.72***	.40***	.49***	.33***	.50***
PPGM-I				.70***	.37***	.41***	.31***	.37***
PPGM-O					.34***	.41***	.28***	.40***
GBS-AY						.79***	.91***	.72***
GBS-AM							.63***	.90***
GBS-FY								.70***

Notes: ***p<.001; PGSI = Problem Gambling Severity Index score; PPGM = Problem and Pathological Gambling Measure; PPGM-T = total score; PPGM-P = problems score; PPGM-I = impaired control score; PPGM-O = other issues score; GBS = Gambling Behavior Scale; GBS-AY = diversity of gambling activities in the last year; GBS-AM = diversity of gambling activities in the last month; GBS-FY = gambling frequency in the last year; GBS-FM = gambling frequency in the last month.

Table 5. Correlations between gambling measures and relevant individual differences and behaviour.

	BSCS	EIS	BSSS	CT	PRDS	DOS-T
PGSI	−.39***	.39***	.20***	.12**	.12**	.40***
PPGM-T	−.32***	.37***	.17***	.12**	.11**	.32***
PPGM-P	−.32***	.31***	.15***	.09**	.10**	.29***
PPGM-IC	−.25***	.37***	.16***	.14***	.10**	.29***
PPGM-O	−.27***	.32***	.14***	.09*	.08*	.28***
GBS-A-Y	−.25***	.28***	.22***	.10**	−.03	.27***
GBS-A-M	−.20***	.23***	.13***	.09*	−.04	.25***
GBS-F-Y	−.22***	.26***	.24***	.10*	−.05	.28***
GBS-F-M	−.20***	.24***	.15***	.06	−.06	.28***

Notes: *p<.05; **p<.01; ***p<.001; PGSI = Problem Gambling Severity Index score; PPGM = Problem and Pathological Gambling Measure; PPGM-T = total score; PPGM-P = problems score; PPGM-I = impaired control score; PPGM-O = other issues score; GBS = Gambling Behavior Scale; GBS-AY = diversity of gambling activities in the last year; GBS-AM = diversity of gambling activities in the last month; GBS-FY = gambling frequency in the last year; GBS-FM = gambling frequency in the last month. BSCS = self-control; EIS = impulsivity; BSSS = sensation-seeking; CT = choice task; PRDS = relative deprivation; DOS-T = risk attitudes total score.

Table 6. Correlations between gambling measures and domain-specific risk attitudes.

	DOS-S	DOS-R	DOS-G	DOS-I	DOS-HS	DOS-E
PGSI	.01	.25***	.48***	.22***	.34***	.42***
PPGM-T	−.01	.19***	.42***	.16***	.27***	.36***
PPGM-P	−.06	.18***	.37***	.13***	.25***	.35***
PPGM-IC	.05	.17***	.36***	.15***	.24***	.30***
PPGM-O	.00	.13***	.39***	.16***	.23***	.31***
GBS-A-Y	−.07	.17***	.38***	.20***	.22***	.28***
GBS-A-M	−.08*	.15***	.37***	.20***	.19***	.28***
GBS-F-Y	−.02	.18***	.35***	.21***	.24***	.25***
GBS-F-M	−.06	.19***	.37***	.21***	.22***	.28***

Notes:*p<.05; ***p<.001.; PGSI = problem gambling tendencies; DOS-S = social risk attitudes; DOS-R = recreational risk attitudes; DOS-G = gambling risk attitudes; DOS-I = investment risk attitudes; DOS-HS = health/safety risk attitudes; DOS-E = ethical risk attitudes.

21 unreported gender; age: M = 37.0, SD = 12.9, Range: 18 to 77) completed wave one. There were 363 participants (192 women, 168 men, 3 trans*, 1 unreported gender; age: M = 37.7, SD = 13.2, range: 18 to 77) who provided an email address allowing contact for the second-wave of recruitment. A total of 222 participants (121 women, 100 men, 2 trans,* 3 unreported gender; age: M = 39.3, SD = 13.4, range: 18 to 78) returned to complete the second wave (a return rate of 61.1%).[1]

Participants completed basic demographic measures: age, gender, highest educational attainment, current relationship status and job status. Problem gambling tendencies were measured using the PGSI and the PPGM. Participants also completed an attention check as described in Goodman et al. (2013).

Results and discussion

Data preparation

Among the 222 participants who completed both survey waves, 200 (90.1%) successfully completed the attention check in both. Participants who failed the attention check were not included in the results reported below. Descriptive statistics for these participants (employment status, highest level of education, and relationship status) are provided in Table 1.

Gambling: wave one

PGSI. The reliability of the PGSI measure was high (*Cronbach's α* = .92). Descriptive statistics for the PGSI were as follows: *M* = 1.88; *SD* = 3.69. Participants were categorized as follows: non-gamblers or non-problem gamblers (57.1%); at-risk gamblers (29.3%); problem gamblers (13.6%).

PPGM. The reliability of the PPGM measure was acceptably high (*Cronbach's α* = .88). Descriptive statistics for the PPGM were as follows: total score: *M* = 1.05, *SD* = 2.14; problems subscale: *M* = .29; *SD* = .95; impaired control subscale: *M* = .56; *SD* = .92; other issues subscale: *M* = .21, *SD* = .58. Participants were categorized as follows: non-gamblers (62.5%); recreational gamblers (18.0%); at-risk gamblers (7.5%); problem gamblers (4.0%); pathological gamblers (8.0%).

Gambling: wave two

PGSI. The reliability of the PGSI measure was high (*Cronbach's α* = .92). Descriptive statistics for the PGSI were as follows: *M* = 1.83; *SD* = 3.63. Participants were categorized as follows: non-problem gamblers or non-gamblers (58.5%); at-risk gamblers (27.0%); problem gamblers (14.5%).

PPGM. The reliability of the PPGM measure was acceptably high (*Cronbach's α* = .86). Descriptive statistics for the PPGM were as follows: total score: *M* = .97, *SD* = 1.89; problems subscale: *M* = .23; *SD* = .67; impaired control subscale: *M* = .53; *SD* = .93; other issues subscale: *M* = .21, *SD* = .60. Participants were categorized as follows: non-gamblers (66.0%); recreational gamblers (13.5%); at-risk gamblers (9.0%); problem gamblers (6.0%); pathological gamblers (5.5%).

Gambling test-retest reliability

Test-retest reliability of the PGSI was very good (*r* = .87). Participants' problem gambling severity was categorized identically by PGSI scores (as non-problem gamblers or non-gamblers, at-risk, or problem/pathological gambling) in 83.3% of cases from wave one to wave two (Cohen's κ = .70, $p<.001$). A McNemar-Bowker test indicated that the proportion of categorizations did not significantly differ between the two waves ($p = .28$). Interestingly, the three-to-four-week test-retest reliability observed in the present study was higher than previously reported three-to-four-week test-retest reliability of the PGSI (.78; Ferris & Wynne, 2001).

Test-rest reliability of PPGM total scores was somewhat lower (*r* = .66), but was comparable to some previous test-retest estimates of mental health measures on crowdsourced platforms (Shapiro et al., 2013). Participants' problem gambling severity was categorized identically by the PPGM (as non-problem, recreational, at-risk, problem or pathological) in 69.5% of cases from wave one to wave two (Cohen's κ = .25, $p<.001$). A McNemar-Bowker test indicated that the proportion of categorizations significantly differed across the two waves ($p < .001$). Previous research produced a test-retest reliability of .78 for the PPGM total score (Williams & Volberg, 2014).

Our results suggest that reliability of the PGSI measure on Crowdflower is high. Our results also suggest that the psychometrics of the PGSI appear to be superior to those of the PPGM in the Crowdflower sample used. Both internal validity and test-retest reliability

was substantially higher in the PGSI than in the PPGM; in fact, PGSI validity and reliability in our sample were higher than previously demonstrated in other samples (e.g. Williams & Volberg, 2014). We again observed rates of problem gambling on Crowdflower that are substantially higher than in the general population.

General discussion

Results from four studies indicated that two popular crowdsourcing platforms – Crowdflower and Amazon Mechanical Turk – include an extraordinary proportion of self-reported problem gamblers. Problem gamblers (categorized using the benchmark PGSI) comprised 24.5%, 21.9%, 9.0%, and 13.6 to 14.5% of participants in Studies 1, 2, 3, and 4, respectively. Relative to an average past-year population-level problem gambling prevalence of 2.3% (Williams et al., 2012), the current samples have significantly higher rates. In Studies 2 and 3, problem gambling tendencies were consistently associated as expected with individual differences in personality, risk attitudes, affect, and behavioural risk-taking. Study 3 evidenced general gambling involvement as strongly associated with all expected outcome measures (with the exceptions of behavioural risk-taking and relative deprivation). Together, the results suggest that crowdsourcing platforms offer access to samples with remarkably high rates of gambling and problem gambling tendencies.

Crowdsourced sample representativeness

Demographic characteristics of our samples were similar to population parameters in the USA and Canada (the two most common sources of crowdsourced participants). The proportion of women in our samples ranged from 49.4% to 53.1%, which accords with population estimates in the USA (50.8%; US Census Bureau, 2011) and Canada (50.4%; Statistics Canada, 2011b). Mean age in our samples ranged from 34.2 to 39.3, similar to estimates from the USA (37.9 years) and Canada (42 years) (CIA, 2016). The proportion of people employed full-time in our samples ranged from 45.1% to 54.5%; population estimates are 56.8% in Canada (Statistics Canada, 2015) and 52.3% in the USA (Bureau of Labor Statistics, 2016). In the two samples where we collected relationship status, 38.5% and 51.0% reported being currently married (or equivalent); population estimates are 47.8% in Canada (Statistics Canada, 2016) and 50% in the USA (US Census Bureau, 2016). In the one sample where we collected ethnicity data, 78.8% of people reported being Caucasian. The proportion of Caucasian people is 76.7% in Canada (Statistics Canada, 2011a) and 72.4% in the USA (US Census Bureau, 2016). All the demographic comparisons are approximations owing to different sampling methods. However, the general observation stands: demographic characteristics from the Crowdflower and MTurk samples were collectively remarkably similar to population estimates.

Despite the demographic similarities between the present crowdsourced samples and population parameters, reports of gambling behaviours and problem gambling tendencies were remarkably high in our samples. Consequently, it appears that crowdsourced samples are not representative of more general populations with respect to gambling behaviours and problem gambling tendencies. Cross-national surveys have indicated that past-year problem gambling prevalence estimates range from 0.5% to 7.6%, with an average of 2.3% (Williams et al., 2012). Our problem gambling estimates

for crowdsourced samples ranged from 9.0% to 24.5%, a much higher estimate. These estimates are also much higher than those observed from studies utilizing market panels (e.g. Williams et al., 2012), as well as undergraduate student samples (e.g. Gainsbury, Russell, & Blaszczynski, 2014).

Although gambling behaviours and problem tendencies were relatively high in our samples, the magnitude of associations between problem gambling and general gambling involvement and relevant traits, attitudes and behaviours were remarkably similar to those obtained in previous work with both student and community samples (Mishra, Lalumière, Morgan, & Williams, 2010; Mishra, Lalumière, & Williams, 2010, in press). As examples, the following are correlation magnitude associations between general gambling involvement, problem gambling, and various measures obtained in the present study compared to results obtained in Mishra et al. (2010) among a convenience sample of undergraduates: self-control (present: r = -.20 to -.39; previous: r = -.24 and .39); impulsivity (present: r = .23 to .39; previous: r = .21 and .32); sensation-seeking (present: r = .13 to .24; previous: r = .24 and .30); behavioural risk-taking in the Choice Task (present: r = .06 to .14; previous: r = .11 and .13); overall risk-attitudes (present: r = .25 to .40; previous: r = .36 and .37).

The proportion of Crowdflower participants who met criteria for problem gambling (in Studies 1, 2, and 4) was higher than the proportion of problem gamblers in the MTurk sample (Study 3). On MTurk, participants are directly paid in cash; however, Crowdflower utilizes multiple channel partners, each of which pays participants in different ways. Some use cash, whereas others use raffle tickets or some other proprietary site-specific currency, which may attract a broader array of participants predisposed to gambling.

Strengths and limitations

The current studies have several strengths and limitations that provide direction for future research. All four samples were very large and were collected at low overall cost. Accordingly, crowdsourced samples offer opportunities for more regular reporting of high statistical power studies.

Participants may have completed more than one of the studies reported in the article. Consequently, portions of our four samples may be non-independent. The impossibility of ensuring sample independence is a limitation for all multi-study papers (and even replications across papers) when participant data is collected anonymously. The extent and impact of such a limitation are also necessarily unknowable given the methods. Future researchers concerned with such limitations should consider innovating novel designs that allow for anonymity while ensuring absolute sample independence.

Studies 1 and 2 did not include the attention checks provided in Studies 3 and 4. Some previous research suggests that attention checks are important for validity (e.g. Goodman et al., 2013); other research has indicated attention checks are overly intrusive and do not necessarily enhance validity (because participants on crowdsourcing platforms may be overly vigilant or sensitive to attention checks; Chandler & Shapiro, 2016). We note that the patterns of results from Studies 3 and 4 were entirely unchanged regardless of whether or not those who failed attention checks were included in our analyses, magnitudes for descriptive statistics of all measures were nearly identical, and all associations remained

significant (or non-significant). Other passive attention checks have been used in some studies, such as outliers in time-to-completion. We did not utilize such passive attention checks; future studies might benefit from such filters. We note, however, that our very large sample sizes minimize the effect of any outliers on our results.

Being a gambler was not an inclusion criterion for any of our studies, and the PGSI did not facilitate distinguishing gamblers and non-gamblers. Future research should examine whether more specific calls for gamblers in recruitment descriptions lead to even higher proportions of gambling behaviours and problem gambling tendencies in crowdsourced samples.

Not all measures were administered across all four studies. Consequently, direct comparisons of demographics and correlation magnitudes between our Crowdflower samples (Studies 1, 2, and 4) and our MTurk sample (Study 3) are difficult (except in the case of directly overlapping measures and associations). Nevertheless, the patterns of results across all four studies are remarkably consistent.

Conclusion

Crowdsourcing platforms have rapidly increased in popularity among behavioural scientists due to ease of data collection and low cost. The results of our four studies demonstrate that such platforms may be of utility to gambling researchers. Among participants on Crowdflower and Amazon Mechanical Turk, we demonstrated (a) convergent associations across various measures of problem gambling tendencies and general gambling involvement; (b) associations between gambling measures and relevant personality traits, risk attitudes, affect, and behavioural risk-taking with magnitudes very similar to those observed in previous research; and (c) adequate test-retest reliability of problem gambling measures across time. We also observed rates of gambling and problem gambling tendencies that were far higher than in the general population. Although crowdsourced samples appear not to be representative of the general population with regards to gambling behaviour, they may still be useful for recruiting participants who exhibit relatively high levels of gambling and problem gambling tendencies.

Conflicts of interest

The authors have no conflicts of interest to declare.

Competing interests

The funding source had no influence or involvement in any aspects of the research. The authors acknowledge that they have no financial interests and benefits that arise from the direct applications of this research.

Constraints on publishing

The authors acknowledge that this work was not subject to any constraints on publishing.

Funding sources

This research was supported by an Insight Development Grant from the Social Sciences and Humanities Research Council of Canada to SM.

Note

1. We note that participants who did not complete wave two reported significantly higher PGSI scores (t = 2.54, p = .01) and were significantly younger (t = -2.34, p = .02) than those who returned to complete wave two, although the effect sizes of these differences were small (ds of .27 and .25, respectively).

Notes on contributors

Sandeep Mishra is an assistant professor of organizational behavior at the University of Regina. He has published over 40 journal articles and book chapters exploring diverse research questions in the broad areas of decision-making under risk and uncertainty, personality and individual differences, and evolutionary theory.

R. Nicholas Carleton is an associate professor of psychology at the University of Regina. He is actively involved in clinical and experimental research, with his interests including the measurement, assessment and treatment of anxiety- and trauma-related disorders. He has published over 100 peer-reviewed journal articles and book chapters examining wide-ranging research questions in the area of mental health. He enjoys teaching and supervision of undergraduate and graduate students, and maintains a small private practice for military and first responders who have post-traumatic stress disorder.

References

Arthur, D., Tong, W. L., Chen, C. P., Hing, A. Y., Sagara-Rosemeyer, M., Kua, E. H., & Ignacio, J. (2008). The validity and reliability of four measures of gambling behaviour in a sample of Singapore University students. *Journal of Gambling Studies, 24*, 451–462.

Berinsky, A. J., Huber, G. A., & Lenz, G. S. (2012). Evaluation of online labor markets for experimental research: Amazon.com's Mechanical Turk. *Political Analysis, 20*, 351–368.

Beshai, S., Mishra, S., Meadows, T. J. S., Parmar, P., & Huang, V. (2017). Minding the gap: Subjective relative deprivation and depressive symptoms. *Social Science & Medicine, 173*, 18–25.

Blais, A.-R., & Weber, E. U. (2006). A Domain-Specific Risk-Taking (DOSPERT) scale for adult populations. *Judgment and Decision Making, 1*, 33–47.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science, 6*, 3–5.

Bureau of Labor Statistics. (2016). *Labor force statistics from the current population survey*. Retrieved from https://www.bls.gov/cps/cpsaat08.htm

Callan, M. J., Ellard, J. H., Will Shead, N. W., & Hodgins, D. C. (2008). Gambling as a search for justice: Examining the role of personal relative deprivation in gambling urges and gambling behavior. *Personality and Social Psychology Bulletin, 34*, 1514–1529.

Callan, M. J., Will Shead, N. W., & Olson, J. M. (2015). The relation between personal relative deprivation and the urge to gamble among gamblers is moderated by problem gambling severity: A meta-analysis. *Addictive Behaviors, 45*, 146–149.

Chandler, S. & Shapiro, D. (2016). Conducting clinical research using crowdsourced convenience samples. *Annual Review of Clinical Psychology, 12*, 53–81.

CIA. (2016). *The world factbook*. Retrieved from https://www.cia.gov/library/publications/the-world-factbook/fields/2177.html

Colasante, E., Gori, M., Bastiani, L., Siciliano, V., Giordani, P., Grassi, M., & Molinaro, S. (2013). An assessment of the psychometric properties of italian version of CPGI. *Journal of Gambling Studies, 29*, 765–774.

Crump, M. J. C., McDonnell, J. V., & Gureckis, T. M. (2013). Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research. *PLoS ONE, 8*, e57410.

Eysenck, S. B. G., Pearson, P. R., Easting, G., & Allsopp, J. F. (1985). Age norms for impulsiveness, venturesomeness and empathy in adults. *Personality and Individual Differences, 6*, 613–619.

Ferris, J., & Wynne, H. (2001). *The Canadian problem gambling index: Final report* (Submitted to the Canadian Centre on Substance Abuse). Ottawa, Ontario: Canadian Centre on Substance Abuse.

Fessler, D. M. T., Pillsworth, E. G., & Flamson, T. J. (2004). Angry men and disgusted women: An evolutionary approach to the influence of emotions on risk-taking. *Organizational Behavior and Human Decision Processes, 95*, 107–123.

Fort, K., Adda, G., & Cohen, K. B. (2011). Amazon Mechanical Turk: Gold mine or coal mine? *Computational Linguistics, 37*, 413–420.

Gainsbury, S. M., Russell, A., & Blaszczynski, A. (2014). Are psychology university student gamblers representative of non-university students and general gamblers? A comparative analysis. *Journal of Gambling Studies, 30*, 11–25.

Goodman, J. K., Cryder, C. E., & Cheema, A. (2013). Data collection in a flat world: The strengths and weaknesses of Mechanical Turk samples. *Journal of Behavioral Decision Making, 26*, 213–224.

Hollingshead, S. J., Kim, H. S., Wohl, M. J. A., & Derevensky, J. L. (2016). The social casino gambling-gambling link: Motivation for playing social casino games determines whether self-reported gambling increases or decreases among disordered gamblers. *Journal of Gambling Issues, 33*, 52–67.

Hoon, A. E., & Dymond, S. (2013). Altering preferences for concurrently available simulated slot machines: Nonarbitrary contextual control over gambling choice. *Analysis of Gambling Behavior, 7*, 35–52.

Horton, J. J., Rand, D. G., & Zeckhauser, R. J. (2011). The online laboratory: Conducting experiments in a real labor market. *Experimental Economics, 14*, 399–425.

Hoyle, R. H., Stephenson, M. T., Palmgreen, P., Lorch, E. P., & Donohew, R. L. (2002). Reliability and validity of a brief measure of sensation seeking. *Personality and Individual Differences, 32*, 401–414.

Kim, H. S., & Hodgins, D. C. (in press). Reliability and validity of data obtained from alcohol, cannabis, and gambling populations on Amazon's Mechanical Turk. *Psychology of Addictive Behaviors*.

Kim, H. S., Wohl, M. J. A., Salmon, M. M., Gupta, R., & Derevensky, J. (2015). Do social casino gamers migrate to online gambling? An assessment of migration rate and potential predictors. *Journal of Gambling Studies, 31*, 1819–1831.

Loo, J. M. Y., Oei, T. P., & Raylu, N. (2011). Problem gambling, gambling correlates, and help-seeking attitudes in a Chinese sample: An empirical evaluation. *Psychology, 2*, 342–354.

Mishra, S., & Lalumière, M. L. (2010). You can't always get what you want: The motivational effect of need on risk-sensitive decision-making. *Journal of Experimental Social Psychology, 46*, 605–611.

Mishra, S., & Novakowski, D. (2016). Personal relative deprivation and risk: An examination of individual differences in personality, attitudes, and behavioral outcomes. *Personality and Individual Differences, 90*, 22–26.

Mishra, S., Lalumière, M. L., & Williams, R. J. (2010). Gambling as a form of risk-taking: Individual differences in personality, behavioral preferences for risk, and risk-accepting attitudes. *Personality and Individual Differences, 49*, 616–621.

Mishra, S., Lalumière, M. L., Morgan, M., & Williams, R. J. (2010). An examination of the relationship between gambling and antisocial behavior. *Journal of Gambling Studies, 27*, 409–426.

Mishra, S., Lalumière, M. L., Morgan, M., & Williams, R. J. (2011). An examination of the relationship between gambling and antisocial behavior. *Journal of Gambling Studies, 27*, 409–426.

Mishra, S., Morgan, M., Lalumière, M. L., & Williams, R. J. (2010). Mood and audience effects on video lottery terminal gambling. *Journal of Gambling Studies, 26*, 373–386.

Mishra, S., Lalumière, M. L., & Williams, R. J. (in press). Gambling, risk-taking, and antisocial behavior: A replication study supporting the generality of deviance. *Journal of Gambling 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.
Paolacci, G., & Chandler, J. (2014). Inside the turk: Understanding Mechanical Turk as a participant pool. *Current Directions in Psychological Science, 23*, 184–188.
Paolocci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making, 5*, 411–419.
Raylu, N., & Oei, T. P. S. (2002). Pathological gambling: A comprehensive review. *Clinical Psychology Review, 22*, 1009–1061.
Shapiro, D. N., Chandler, J., & Mueller, P. A. (2013). Using Mechanical Turk to study clinical populations. *Clinical Psychological Science, 1*, 213–220.
Sprouse, J. (2011). A validation of Amazon Mechanical Turk for the collection of acceptability judgments in linguistic theory. *Behavior Research Methods, 43*, 155–167.
Statistics Canada. (2011a). *National household survey, 2011*. Retrieved from https://www12.statcan.gc.ca/nhs-enm/2011/dp-pd/dt-td/index-eng.cfm
Statistics Canada. (2011b). *Women in Canada: A gender-based statistical report* (Catalogue no. 89-503-X). Retrieved from http://www.statcan.gc.ca/pub/89-503-x/89-503-x2010001-eng.pdf
Statistics Canada. (2015). *Full-time Employment, 1976 to 2014* (Catalogue no. 11-626-049). Retrieved from http://www.statcan.gc.ca/pub/11-626-x/11-626-x2015049-eng.htm
Statistics Canada. (2016). *Population by marital status and sex*. Retrieved from http://www.statcan.gc.ca/tables-tableaux/sum-som/l01/cst01/famil01-eng.htm
Stinchfield, R., McCready, J., Turner, N. E., Jiminez-Murcia, S., Petry, N. M., Grant, J., … Winters, K. C. (2016). Reliability, validity, and classification accuracy of the DSM-5 diagnostic criteria for gambling disorder and comparison to DSM-IV. *Journal of Gambling Studies, 32*, 905–922.
Tabri, N., Dupuis, D. R., Kim, H. S., & Wohl, M. J. A. (2015). Economic mobility moderates the effect of relative deprivation on financial gambling motives and disordered gambling. *International Gambling Studies, 15*, 309–323.
Tangney, J. P., Baumeister, R. F., & Boone, A. L. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality, 72*, 271–324.
Toneatto, T. & Nguyen, L. (2007). Individual characteristics and problem gambling behavior. In G. Smith, D. C. Hodgins, & R. J. Williams (Eds.), *Research and measurement issues in gambling studies* (pp. 279–303). Amsterdam: Elsevier.
U.S. Census Bureau. (2011). Age and sex composition: 2010. Retrieved from http://www.census.gov/prod/cen2010/briefs/c2010br-03.pdf
U.S. Census Bureau. (2016). *Marital status of the population 15 years old and over, by sex, race, and Hispanic origin: 1950 to present*. Retrieved from https://www.census.gov/hhes/families/data/marital.html
Van Brunschot, E. G. (2009). *Gambling and risk behaviour: A literature review*. Report prepared for the Alberta Gambling Research Institute.
Williams, R. J., & Volberg, R. A. (2010). *Best practices in the population assessment of problem gambling*. Report prepared for the Ontario Problem Gambling Research Centre, Guelph, Ontario, Canada. March 31, 2010.
Williams, R. J., & Volberg, R. A. (2014). Classification accuracy of four problem gambling assessment instruments. *International Gambling Studies, 14*, 15–28.
Williams, R. J., Royston, J., & Hagen, B. F. (2005). Gambling and problem gambling within forensic populations: A review of the literature. *Criminal Justice and Behavior, 32*, 665–689.
Williams, R. J., Volberg, R. A., & Stevens, R. M. G. (2012). *The population prevalence of problem gambling: Methodological influences, standardized rates, jurisdictional differences, and worldwide trends*. Report prepared for the Ontario Problem Gambling Research Centre and the Ontario Ministry of Health and Long-Term Care. May 8, 2012. Retrieved from http://hdl.handle.net/10133/3068