The Ordering of Gambling Severity and Harm Scales: A Cautionary Tale

Kate Sollis,[1] Patrick Leslie,[2] Nicholas Biddle,[1] & Marisa Paterson[3]

[1] Centre for Social Research and Methods, Australian National University
[2] School of Politics and International Relations, Australian National University
[3] Centre for Gambling Research, Australian National University

Abstract

Question-order effects are known to occur in surveys, particularly those that measure subjective experiences. The presence of context effects will impact the comparability of results if questions have not been presented in a consistent manner. In this study, we examined the influence of question order on how people responded to two gambling scales in the Australian Capital Territory Gambling Prevalence Survey: The Problem Gambling Severity Index and the Short Gambling Harm Screen. The application of these scales in gambling surveys is continuing to grow, the results being compared across time and between jurisdictions, countries, and populations. Here we outline a survey experiment that randomized the question ordering of these two scales. The results show that question-order effects are present for these scales, demonstrating that results from them may not be comparable across jurisdictions if the scales have not been presented consistently across surveys. These findings highlight the importance of testing for the presence of question-order effects, particularly for those scales that measure subjective experiences, and correcting for such effects where they exist by randomizing scale order.

Keywords: PGSI, SGHS, problem gambling, question-order effects, context effects, gambling prevalence, survey experiment

---

Introduction

A range of factors can influence the way an individual responds to questions in a survey. Aside from the content itself, a response can be affected by the topic of interest, the location of enumeration, the phrasing of questions, and the response format (Deaton & Stone, 2016; OECD, 2013; Sudman et al., 1996). These

phenomena are more broadly known as context effects, a form of measurement error in the total survey error framework (Groves & Lyberg, 2010). One kind of context effect is the ordering of questions in surveys, known as question-order effects.

A large body of literature illustrates the impact of question ordering on how people respond in surveys (Deaton & Stone, 2016; Schwarz, 1999; Stark et al., 2018; Strack et al., 1991; Sudman et al., 1996; Tourangeau & Rasinski, 1988). Such studies have examined question-order effects in surveys that measure attitudes (Lasorsa, 2003; McFarland, 1981; Moore, 2002; Stark et al., 2018; Tourangeau & Rasinski, 1988; Tourangeau et al., 2003), subjective well-being, and life satisfaction or evaluation (Deaton & Stone, 2016; Garbarski et al., 2015; Lee et al., 2016; McClendon & O'Brien, 1988); these effects have also been examined in Delphi studies (Brookes et al., 2018).

Subjective measures are generally found to be particularly sensitive to context effects, such as when individuals construct their answers on being presented with a question, as opposed to retrieving an answer from their memory, as would be done for more objective questions (Sudman et al., 1996). Deaton and Stone (2016) also suggest that questions that are difficult to answer, such as subjective well-being, are more sensitive to context effects. A number of studies have empirically tested the influence of context effects for subjective measures. Although there is some variation in the question topics and results, a common theme among them is that question-order effects are higher for those who provide a negative response to the preceding question. That is, question-order effects do not appear to be consistent across population subgroups within a sample.

For example, Deaton and Stone (2016) examined how questions regarding satisfaction with politics influenced questions on life evaluation, finding that there was a subtractive effect for life evaluation (meaning that life evaluation was worse when this question followed satisfaction with politics), but only for those who disapproved of the way the country was going. Garbarski et al. (2015) explored how self-rated health is affected by health questions in specific domains, finding a subtractive and assimilation effect, meaning that, on average, people reported lower self-rated health that more closely aligned with their reports on domain-specific health when this question came after domain-specific health questions. This result was driven by people with a large number of health risks. In assessing the impact of question ordering between life-satisfaction and self-rated health questions, Lee et al. (2016) found an assimilation effect for life satisfaction, with the correlation increasing when life satisfaction came after the question on self-rated health. This effect was even stronger for those with chronic conditions, leading the authors to recommend these measures being placed apart in surveys.

Gambling surveys tend to include subjective questions that ask respondents to reflect on their own gambling behaviour and, as a result, may be highly prone to question-order effects. Previous research on question-order effects in gambling surveys confirms this possibility, with one study examining the impact of survey placement in

the Focal Adult Gambling Screen (Harrison et al., 2017) and another the placement of simple lottery questions and perceptions of risk propensity (Golik, 2019), both studies finding question-order effects. This finding highlights the importance of testing for question-order effects in gambling surveys, particularly those that are of a subjective nature.

Previous literature suggests that question-order effects are frequently observed for subjective measures. Although the research base is limited, it indicates that question-order effects are prevalent in gambling surveys. In the present study, we sought to contribute to this literature base by examining question-order effects in the Problem Gambling Severity Index (PGSI) and the Short Gambling Harms Scale (SGHS), two gambling scales for which question-order effects have not yet been tested. Such research will provide methodological advances in the social sciences by providing a deeper understanding about the types of questions prone to order effects, as well as in the gambling field more specifically by offering guidance on the placement of these two scales in a survey.

Method

Data

In this study, we used data from the 2019 Australian Capital Territory (ACT) Gambling Survey, commissioned by the ACT Gambling and Racing Commission. A total of 10,000 ACT residents aged 18 years or over were interviewed in the survey throughout a 6-week period (April to May 2019). Participants provided detailed information on their gambling participation, expenditure, and frequency. They also answered two sets of subjective questions on problem gambling (PGSI) and gambling harms (SGHS). These questions are discussed in detail in the following sections, and further information on the survey is outlined by Paterson et al. (2019).

Sample Characteristics

A dual-frame sample design was used that consisted of randomly generated (random digit dialling) landline telephone numbers and listed mobile phone numbers. To address the gradual decline in the population's use of landline telephones, the survey authors set a ratio of 70% mobile to 30% landline numbers to improve the population coverage of the survey over previous landline-only surveys. The overall response rate for the survey was 16.3%. Table 1, taken from Paterson et al. (2019), outlines the socio-demographic characteristics of the respondents, as well as the corresponding figures for the ACT population used for benchmarking.

The PGSI and SGHS Scales

Gamblers were provided with the PGSI and the SGHS scales, which are designed to screen for problem gambling and gambling harm, respectively. In 2001, the PGSI

Table 1

Sample characteristics (unweighted) of respondents

Characteristic	Total, %	ACT adults, %
Total	$n = 10,000$	$N = 259$
Gender		
Male	46.4	48.3
Female	53.4	51.1
Age, years		
18–24	4.2	14.1
25–34	13.8	21.5
35–44	18.8	18.9
45–54	19.2	16.1
55–64	18.4	13.1
$\geqslant$65	25.7	16.4
Marital status		
Married	57.4	^
De facto/in a relationship	12.1	^
Separated and single	3.1	^
Divorced and single	6.7	^
Widowed and single	5.2	^
Single	15.2	^
Born in Australia		
Yes	73.8	66.8
Education		
University	54.6	40.5
TAFE certificate/diploma	21.5	^
Completed year 12	13.9	^
Completed year 11	1.1	^
Completed year 10	4.5	^
Completed years 7–9	1.2	^
Completed primary school	0.2	^
Other	2.3	^
Work status		
Employed full time	50.9	^
Employed part time or casual	16.7	^
Unemployed, looking for work	1.7	^
Not in the paid labour force	30.2	^

Note. ACT = Australian Capital Territory; TAFE = Technical and Further Education. ^ indicates that this variable was not used to benchmark the target ACT population.

ACT population statistics sourced from the Australian Bureau of Statistics Census 2016. Adapted from *2019 ACT Gambling Survey*, by M. Paterson, P. Leslie, and M. Taylor, 2019 (https://csrm.cass.anu.edu.au/sites/default/files/docs/2019/10/2019-ACT-Gambling-Survey.pdf).

was developed in response to calls for an appropriate and validated measure to identify pathological or problem gambling—based on clinical criteria—in general population surveys. The PGSI has become the primary measure for establishing the prevalence of problem gambling both internationally and in Australia. It differs from an individual diagnostic or clinical tool in that it does not explicitly measure gambling harm, but rather a mixture of pathological gambling symptoms; external

indicators of problem gambling; and negative consequences for the gambler, their social network, or the community (Ferris & Wynne, 2001).

Distinct from the PGSI, the SGHS was developed to directly measure the harm experienced by gamblers (Browne et al., 2018) and is adapted from a 72-item scale that compiles a more comprehensive list of gambling-related harm (Langham et al., 2016). The shortened scale asks 10 questions about gambling harms and whether respondents had experienced these harms in the past 12 months. The number of harms reported are totalled to give the individual an SGHS score of between 0 and 10.

The PGSI and SGHS items were asked only of respondents who had gambled in the last 12 months. The scales were placed together toward the end of the survey, following questions on demographic information, gambling participation, attitudes toward gambling, past gambling behaviour, and self-identification of gambling problems.

Methodology

We tested the impact of the placement order of the PGSI and SGHS in the context of the 2019 ACT Gambling Prevalence Survey. First, we tested for balance in the random assignment of survey participants to the questionnaire ordering, showing that no significant differences in observed socio-demographic indicators appeared to confound the comparison. Second, we performed a simple test of order difference on the combined scores of each scale. Third, we examined the effect of scale order manipulation on the individual items of each scale to reveal which questions were driving assimilation and contrast effects and which were driving additive and subtractive effects. In the following sections, we describe the tests applied and the rationale for the choice of analysis.

Balancing Comparison Groups on Observable Characteristics

We refer to the randomized survey groups as Group A (those who took the PGSI before the SGHS) and Group B (those who took the SGHS before the PGSI). An important step in the analysis, before comparing any differences that questionnaire ordering may have had, was to compare the socio-demographic profiles of the two groups. This step was taken to ensure that the randomization process was effective in producing two samples that were alike in terms of their observable characteristics. We compared Group A and B for age, gender, country of birth, geographic location, education, relationship status, and work status.

Testing for Order Effects

To compare order effects on combined scales and individual scale items, we tested for differences in the combined numeric score and the category threshold. For the numeric score of the PGSI (0–27), this meant comparing the numeric totals of the scale for the 50% who answered the PGSI before the SGHS (Group A) with those of

the 50% who answered after the SGHS (Group B). Similarly, the numeric scores of the SGHS (0–10) were compared based on order. Question-order effects tend to be gauged by using two approaches: additive and subtractive effects, and assimilation and contrast effects (Deaton & Stone, 2016; Garbarski et al., 2015; Lasorsa, 2003; Lee et al., 2016; McFarland, 1981). Additive and subtractive effects measure the extent to which the average response to a question increases (additive) or decreases (subtractive) in its comparative context versus its non-comparative context. The comparative context is the ordering in which the scale or question of interest is asked following questions of interest that may have the potential to cause a priming effect. The non-comparative context is when the ordering is asked prior to any such potentially priming questions of interest. A non-parametric Wilcoxon rank-sum test is used to test for differences.

Assimilation and contrast effects can be defined as the way in which the correlation between two questions changes because of question ordering. If the correlation increases in the comparative context, an assimilation effect is said to occur, as the individual's response becomes more aligned with the previous question. If the correlation decreases, this is referred to as a contrast effect. A multivariate correlation test was conducted to ascertain assimilation and contrast effects. Notably, the presence of order effects does not necessarily inform us about which question placement provides a more accurate response for a given scale. That is, we cannot be certain whether it is the non-comparative or comparative context that is biased. It is feasible that a priming question either influenced participants' responses away from the "truth," or enabled the respondent to think more concretely about a question in its comparative context. The presence of order effects does, however, high-light potential concerns when prevalence rates are compared across different surveys in which the placement of these questions has not been consistent. Figure 1 provides

Figure 1
How Contrast and Assimilation Effects Are Determined for Item 1 of the Short Gambling Harm Screen (SGHS) and Item 9 of the Problem Gambling Severity Index (PGSI) Item 9.
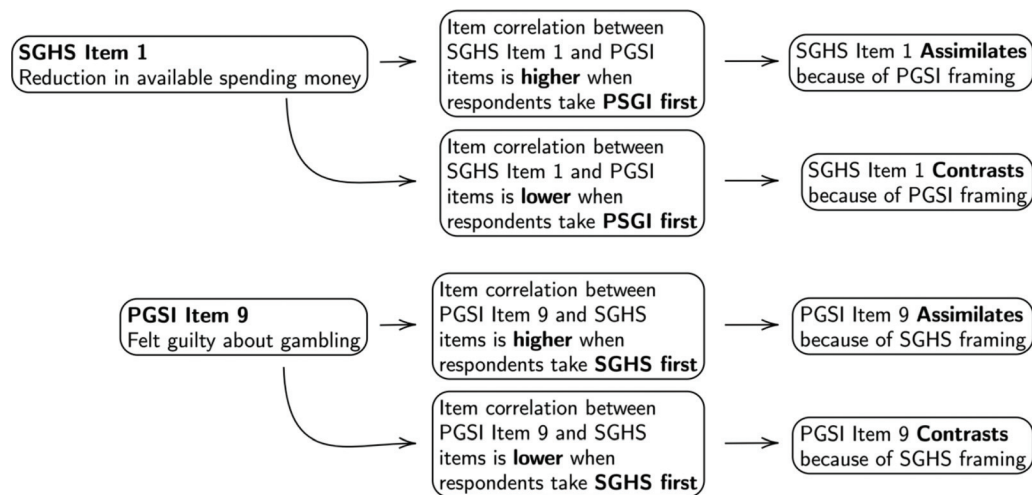
an example of how assimilation and contrast effects are defined for individual scale items.

Results

Balancing Socio-Demographic Characteristics

Results of the Pearson's chi-square test illustrated that the socio-demographic characteristics between Group A and Group B were balanced, as shown in Table 2. This finding demonstrates that the randomization technique resulted in two groups that were similar for observed characteristics.

Effects for Aggregated Scales

The additive and subtractive order effects for the combined scales are shown in Table 3. The question ordering resulted in a significantly different distribution of responses between Group A and Group B for both the PGSI and SGHS scales, with a small effect size. This result is illustrated through the Wilcoxon rank-sum test, where the difference in the PGSI score between the two groups was significant at the 1% significance level with a p-value of .0024. This test also showed a significant difference in the SGHS score between the two groups at the 1% significance level with a p-value of .0083. The rank sums indicate that participants were more likely to respond negatively to the scale that appeared first (PGSI for Group A and SGHS for Group B). This results in a "subtractive" effect, with a small effect size, for both scales.

To better understand what may be contributing to this subtractive effect, we performed a chi-square test for the four PGSI categories that are outlined and validated in Currie et al. (2013) as non-problem gamblers, low-risk, medium-risk, and problem gamblers. We also did so for those who reported any harms through the SGHS scale, as shown in Table 3. For the PGSI scale, these tests illustrated that the subtractive effect observed was driven by those who reported non-problem, low-risk, and medium-risk gambling. Thus, for these groups only, the individuals who completed the SGHS prior to the PGSI reported less risky gambling behaviour. As shown in Table 3, the ordering of the scales affected the prevalence rate by up to 3 percentage points.

The assimilation and contrast order effects are shown in Table 4. First, this table shows that the correlation between the combined PGSI and SGHS scales changed significantly, at the 1% significance level, when the order of the two scales changed, with a small effect size. When PGSI appeared first (Group A), the correlation was .77, compared with a correlation of .66 when SGHS appeared first (Group B). This corresponds to a contrast effect for PGSI and an assimilation effect for SGHS. Thus, members of Group A adapted their responses to the PGSI to *contrast* with the SGHS, and members of Group B adapted their responses to the SGHS to *assimilate* with the PGSI. There are a number of reasons as to why this may happen, and with the given

Table 2

Socio-Demographic Balance of Comparison Groups A and B Assessed From Pearson's χ^2 *Tests*

Variable	Order A, % (PGSI first)	Order B, % (SGHS first)
Gender (*p* = .637)		
Male	49.9	49.2
Female	50.1	50.8
Birthplace (*p* = .305)		
Australia	77.5	76.4
Other	22.5	23.7
Education (*p* = .971)		
Bachelor's degree	26.2	26.4
Certificate/diploma	24.3	23.9
Other	2.1	2.1
Postgraduate degree	22.2	22.6
Year 10	8.6	8.1
Year 12	16.6	17.0
Age, years (*p* = .267)		
18–24	4.3	3.8
25–29	5.6	5.7
30–34	8.4	8.0
35–39	10.3	9.2
40–44	9.5	9.0
45–49	9.3	10.6
50–54	10.1	11.4
55–59	9.7	10.7
60–64	9.6	8.8
65–69	8.8	9.4
≥ 70	14.5	13.5
Relationship (*p* = .398)		
Married	56.3	57.8
De facto/in a relationship	13.6	13.3
Separated	3.3	3.3
Divorced	7.0	6.8
Widowed	4.9	3.8
Single	14.9	15.1
Work status (*p* = .181)		
Employed full time	55.8	57.3
Employed part time and looking for work	1.9	1.6
Employed part time or casual	13.4	13.5
Unemployed, looking for work	1.7	1.1
Not in the paid labour force	27.2	26.5

Note. PGSI = Problem Gambling Severity Index; SGHS = Short Gambling Harms Scale. The total sample for groups A and B combined was 5,851 for PGSI respondents (Group A: 2,932; Group B: 2,919) and 5,886 for SGHS respondents (Group A: 2,952; Group B: 2,934). The total sample size for the Pearson's χ^2 tests varied by socio-demographic category: gender (5,886), age (5,886), birthplace (5,886), education (5,850), relationship status (5,862), and work status (5,861).

data we can only speculate, but one possible explanation is that the two scales are conceptually similar (Tourangeau et al., 2003). When we examined the PGSI groupings individually, no significant contrast or assimilation effect was observed.

Table 3

Additive/Subtractive Effects for PGSI and SGHS Results

Variable/grouping	Order A (PGSI first)	Order B (SGHS first)	Order effect	Effect size
Wilcoxon rank-sum test	Rank sum	Rank sum		
PGSI score (0–27) (Prob > |z| = 0.002)	8696769.5	8423256.5	Subtractive	$r = .040$
SGHS score (0–10) (Prob > |z| = 0.008)	8587313	8738128	Subtractive	$r = -.034$
Pearson's χ^2 test	Percentage	Percentage		
PGSI: non-problem gambling ($p = .002$)	84.8% [83.5%, 86.1%]	87.5% [86.3%, 88.7%]		$V = .040$
PGSI: low risk ($p = .016$)	10.6% [9.5%, 11.8%]	8.8% [7.7%, 9.8%]		$V = -.032$
PGSI: medium risk ($p = .043$)	3.5% [2.8%, 4.2%]	2.6% [2.0%, 3.2%]		$V = -.026$
PGSI: problem gambling ($p = .886$)	1.1% [0.7%, 1.5%]	1.1% [0.7%, 1.5%]		$V = .002$
SGHS > 0 ($p = .006$)	12.2% [11.0%, 13.4%]	14.7% [13.4%, 15.9%]		$V = .036$

Note. PGSI = Problem Gambling Severity Index; SGHS = Short Gambling Harms Scale. 95% confidence intervals in brackets. $V = \sqrt{(\chi^2/N)}$.

Table 4

Contrast/Assimilation Effects for PGSI and SGHS Results

PGSI grouping	Order A (PGSI first)	Order B (SGHS first)	Order effect for PGSI	Order effect for SGHS	Effect size
Effects based on all PGSI groups ($p = .0000$)	0.77	0.66	Contrast	Assimilation	$V = .113$
Effects based on PGSI low-risk group only ($p = .3729$)	0.21	0.28	Nil	Nil	$V = .026$
Effects based on PGSI medium-risk group only ($p = .7388$)	0.48	0.51	Nil	Nil	$V = .064$
Effects based on PGSI high-risk group ($p = .1377$)	0.35	0.00	Nil	Nil	$V = .046$

Note. PGSI = Problem Gambling Severity Index; SGHS = Short Gambling Harms Scale. Multivariate asymptotic chi-square test for the equality of two correlation matrices used to test for differences. $V = \sqrt{(\chi^2/N)}$.

Effects for Individual Scale Items

Given the significant additive and subtractive and contrast and assimilation effects observed for the entire PGSI and SGHS scales, it is worth examining which particular scale items were driving the question-order effects. Table 5 shows the results of the Wilcoxon rank-sum test and the percentage of participants who reported a positive response (i.e., possible gambling harm) for each individual PGSI scale item; it also identifies whether there was a nil, subtractive, or additive effect. Similarly, Table 6 shows the results for each SGHS scale items. Although we did not find individual subtractive effects for each individual item of the PGSI (only Item 7—*felt guilty about your gambling*—showed a significant rank-sum test difference), we observed that each item's rank sum from Order A was greater than its equivalent rank sum from Order B. Under a null hypothesis of no order effect for the entire scale, we calculate a probability of nine consecutive subtractive rank-sum effects ($1/2^9 = 1/512$) under the null hypothesis. This yields a p-value of .002 for the rank sum for PGSI order effects (with the SGHS as a foil), replicating the p-value for the total PGSI Wilcoxon rank sum displayed in Table 3.

The results illustrate that for the PGSI, the subtractive effect is driven by an item that asks respondents whether they have felt guilty about the way they gamble or what happens when they gamble, which affects the summary statistic for this item by over 3 percentage points. The overall subtractive effect for the SGHS scale is driven by three items that ask respondents whether they have experienced a reduction in available money because of gambling, have experienced a reduction in savings because of gambling, and have had regrets that made them feel sorry about their gambling.

To further examine the individual scale items that showed subtractive and additive question-order effects, we tested the assimilation and contrast effects to observe the interaction between the individual scale items in the PGSI and SGHS and the respective aggregate scale for the SGHS and PGSI. As illustrated in Table 7, the correlation between the question on guilt in the PGSI and the SGHS aggregate scale show a contrast effect with a small effect size. This is as expected, with the PGSI scale as a whole having a contrast effect with the SGHS scale as shown in Table 4. Thus, individuals reported lower levels of guilt when this question followed the SGHS scale (known through its subtractive effect), which contrasts with how they responded to the SGHS scale (known through its contrast effect).

In examining the three SGHS scale items that showed significant subtractive and additive effects, we can see that each of these items showed assimilation effects. Thus, for each of these questions, individuals reported lower levels of gambling harm (shown through the respective contrast effects) when it followed the PGSI scale, which assimilates with how they responded to the PGSI scale (known through its assimilation effect).

In summary, this analysis has shown that significant question-order effects are observed between the PGSI and SGHS scales. Although the calculated effect sizes

Table 5

Additive/subtractive Effects for Individual PGSI Items

PGSI scale item	Order A (PGSI first)		Order B (SGHS first)		Order effect	Effect size
	% Reporting positive response	Rank sum	% Reporting positive response	Rank sum		
Bet more than could afford to lose (p = .911)	3.7	8679813	3.7	8622091	Nil	r = .001
Needed to gamble with larger amounts to get the same feeling (p = .763)	3.5	8651196	3.3	8603680	Nil	r = .004
Gone back another day to try to win money lost (p = .718)	4.6	8682896	4.4	8607245	Nil	r = .005
Borrowed money or sold anything to get money to gamble (p = .613)	0.6	8678685	0.7	8634986	Nil	r = -.007
Felt you might have a problem with gambling (p = .803)	3.5	8670770	3.4	8607611	Nil	r = .003
People have criticized your betting or told you that you have a gambling problem (p = .633)	3.4	8681073	3.1	8609067	Nil	r = .006
Felt guilty about the way you gamble or what happens when you gamble (p = .000)	8.5	8807791	5.2	8464713	Subtractive	r = .066
Gambling has caused any health problems, including stress or anxiety (p = .176)	2.6	8697407	2.0	8592734	Nil	r = .018
Gambling has caused financial problems for you or household (p = .206)	1.2	8665213	1.6	8642573	Nil	r = -.016

Note. PGSI = Problem Gambling Severity Index; SGHS = Short Gambling Harms Scale. Wilcoxon rank-sum method used to test differences in the numeric distribution of each scale item. Effect is subtractive where rank sum A > rank sum B for PGSI scale items. $r = Z/\sqrt{N}$.

Table 6

Additive/Subtractive Effects for Individual SGHS Items

SGHS scale item	Order A (PGSI first)		Order B (SGHS first)		Order effect	Effect size
	% Reporting positive response	Rank sum	% Reporting positive response	Rank sum		
Reduction in available money because of gambling ($p = .010$)	6.9	8593611	8.7	8673015	Subtractive	$r = -.034$
Reduction of savings because of gambling ($p = .044$)	5.2	8606810	6.4	8659816	Subtractive	$r = -.026$
Less spending on recreational expenses such as eating out, going to movies ($p = .850$)	2.8	8632725	2.9	8598660	Nil	$r = -.002$
Had regrets that made you feel sorry about your gambling ($p = .093$)	4.5	8598094	5.5	8621552	Subtractive	$r = -.022$
Felt ashamed of your gambling ($p = .324$)	2.3	8657837	1.9	8585291	Nil	$r = .013$
Sold personal items because of your gambling ($p = .771$)	0.2	8677368	0.1	8618653	Nil	$r = .004$
Increased credit card debt because of your gambling ($p = .983$)	0.8	8671319	0.8	8618822	Nil	$r = -.000$
Spent less time with people you care about because of your gambling ($p = .137$)	1.7	8682834	1.2	8583792	Nil	$r = .019$
Felt distressed about your gambling ($p = .480$)	1.8	8657744	2.0	8620637	Nil	$r = -.009$
Felt like a failure because of your gambling ($p = .861$)	1.7	8650497	1.6	8592632	Nil	$r = .002$

Note. PGSI = Problem Gambling Severity Index; SGHS = Short Gambling Harms Scale. Wilcoxon rank-sum method used to test differences in the numeric distribution of each scale item. Effect is subtractive where rank sum **B** > rank sum **A** for SGHS scale items. $r = Z/\sqrt{N}$.

Table 7

Assimilation/Contrast Effects for Items Showing Subtractive/Additive Effects (Correlation Coefficient)

Items contrasted	Order A (PGSI first)	Order B (SGHS first)	Order effect	Effect size
SQ5g (guilty) – SGHS (p = .0218)	0.6095	0.5705	Contrast	V = .030
PGSI – P1a (reduction in available spending money) (p = .001)	0.4275	0.3392	Assimilation	V = .052
PGSI – P1b (reduction of savings) (p = .000)	0.5058	0.3804	Assimilation	V = .078
PGSI – P1d (regrets) (p = .000)	0.5822	0.4724	Assimilation	V = .076

Note. PGSI = Problem Gambling Severity Index; SGHS = Short Gambling Harms Scale. Multivariate asymptotic chi-square test for the equality of two correlation matrices used to test for differences. $V = \sqrt{(\chi^2/N)}$

were small, the ordering affected prevalence rates by up to 3 percentage points, which could be substantial when results are compared across different surveys and population groups. At the aggregate level, subtractive effects were observed for both the PGSI and SGHS scales, meaning that respondents reported lower severity when the scale appeared second. Overall, the PGSI showed a contrast effect, meaning that when the scale appeared second, responses contrasted with the SGHS scale. The SGHS showed an assimilation effect, meaning that when the scale appeared second, responses aligned more closely with the PGSI scale. These effects were driven by particular scale items, with the question on guilt driving the subtractive effect for the PGSI and the questions on reduction in available spending money, reduction in available savings, and feeling regrets from gambling driving the subtractive effect for the SGHS scale.

Discussion

This analysis indicates that question-order effects can be observed between the PGSI and SGHS scales when they are placed together. Without supplementary qualitive research, it is difficult to come to a complete understanding as to why these effects occur, or which scale is more accurate. For both scales, a subtractive effect was observed, with responses to the SGHS becoming more closely aligned with the PGSI responses and responses to the PGSI being more divergent from the SGHS responses. The ordering of the two scales resulted in a divergence of almost 3 percentage points for the estimate of problem gambling based on the PGSI and of 2.5 percentage points based on the SGHS.

Research by Lee et al. (2016) can provide some insight into these effects. They argue that in the context of subjective evaluations, assimilation effects occur because of priming of the previous question, and contrast effects occur when the respondent assumes that the question should be answered in a different way from the previous question. This hypothesis is based on the theory of non-redundancy, which suggests that the cognitive processes involved in completing a survey are similar to those of a

conversation, for which conversational norms dictate that a person assumes a different meaning is implied if asked a similar question twice (Grice, 1975).

Despite not being able to reach a clear conclusion about the psychological reasoning behind these effects, in this study we have demonstrated the importance of testing for question-order effects in gambling harm scales and applying methodological techniques to remove any bias created through question placement. As illustrated, when the PGSI and SGHS scales are placed side by side, each is affected by question-order effects, albeit with small effect sizes. Both scales are widely used for gambling research and policy, and since the development of the SGHS, it has been used in the same surveys as the PGSI (e.g., ORC International, 2018). This usage can be problematic when prevalence rates based on the PGSI and SGHS are compared across surveys with different question placements.

Concluding Comments, Recommendations, and Future Work

In this study, we tested question-order effects for the PGSI and SGHS scales by using the ACT Gambling Prevalence Survey. We found that both the PGSI and SGHS scales display subtractive effects: The respondents decreased the severity of their responses to the scale that was placed second. The PGSI showed a contrast effect, meaning that responses became more divergent from the SGHS when it was asked second, whereas the SGHS showed an assimilation effect. The analysis showed that particular scale items influenced these effects, with the question on guilt driving the PGSI effect and questions on available money, available savings, and regrets driving the SGHS effect.

We therefore recommend that for surveys in which these and similar scales are used, techniques be applied to randomize the order of scales. Doing so has two clear benefits: (1) The question-order effects can be tested to confirm whether such scales are prone to question-order effects, and (2) any bias due to question-order effects that may be present is removed. This study has shown that the PGSI and SGHS exhibit question-order effects and, given that subjective scales are particularly prone to such effects, it is likely that similar scales that measure gambling harm and severity may also be biased through question ordering. Where such randomization has not occurred, we would expect that both the PGSI and SGHS items are biased downward when asked second compared with when they are asked first. We recommend that consideration of this bias be incorporated into the analysis and conclusions when these scales are used.

It is well-known that question-order effects can occur in the context of subjective scales, and this study has contributed to this evidence base by testing two frequently used gambling scales. These results indicate the importance of testing for question-order effects in order to produce more reliable estimates of gambling harm and severity. Testing for these effects in different populations and in different survey contexts (online vs. telephone, for example) is warranted in future work, as well as

testing for whether the presence of other questions between the two scales moderates or exacerbates the question-order effects.

References

Brookes, S. T., Chalmers, K. A., Avery, K. N. L., Coulman, K., & Blazeby, J. M. (2018). Impact of question order on prioritisation of outcomes in the development of a core outcome set: A randomised controlled trial. *Trials, 19*(1), 1–11. https://doi.org/10.1186/s13063-017-2405-6

Browne, M., Goodwin, B. C., & Rockloff, M. J. (2018). Validation of the Short Gambling Harm Screen (SGHS): A tool for assessment of harms from gambling. *Journal of Gambling Studies, 34*(2), 499–512. https://doi.org/10.1007/s10899-017-9698-y

Currie, S. R., Hodgins, D. C., & Casey, D. M. (2013). Validity of the Problem Gambling Severity Index interpretive categories. *Journal of Gambling Studies, 29*(2), 311–327. https://doi.org/10.1007/s10899-012-9300-6

Deaton, A., & Stone, A. A. (2016). Understanding context effects for a measure of life evaluation: How responses matter. *Oxford Economic Papers, 68*(4), 861–870. https://doi.org/10.1093/oep/gpw022

Ferris, J. A., & Wynne, H. J. (2001). *The Canadian problem gambling index*. Canadian Centre on Substance Abuse.

Garbarski, D., Schaeffer, N. C., & Dykema, J. (2015). The effects of response option order and question order on self-rated health. *Quality of Life Research, 24*(6), 1443–1453. https://doi.org/10.1007/s11136-014-0861-y

Golik, J. (2019). Testing question order effects of self-perception of risk propensity on simple lottery choices as measures of the actual risk propensity. *Ask: Research and Methods, 27*, 41–59. https://doi.org/10.18061/ask.v27i1.0003

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics 3: Speech acts* (pp. 41–58). Academic Press.

Groves, R. M., & Lyberg, L. (2010). Total survey error: Past, present, and future. *Public Opinion Quarterly, 74*(5), 849–879. https://doi.org/10.1093/poq/nfq065

Harrison, G., Jessen, L., Lau, M., & Ross, D. (2017). Disordered gambling prevalence: Methodological innovations in a general Danish population survey. *Journal of Gambling Studies*, *34*. https://doi.org/10.1007/s10899-017-9707-1

Langham, E., Thorne, H., Browne, M., Donaldson, P., Rose, J., & Rockloff, M. (2016). Understanding gambling related harm: A proposed definition, conceptual

framework, and taxonomy of harms. *BMC Public Health, 16*(1), 80. https://doi.org/10.1186/s12889-016-2747-0

Lasorsa, D. L. (2003). Question-order effects in surveys: The case of political interest, news attention, and knowledge. *Journalism & Mass Communication Quarterly, 80*(3), 499–512. https://doi.org/10.1177/107769900308000302

Lee, S., McClain, C., Webster, N., & Han, S. (2016). Question order sensitivity of subjective well-being measures: Focus on life satisfaction, self-rated health, and subjective life expectancy in survey instruments. *Quality of Life Research, 25*(10), 2497–2510. https://doi.org/10.1007/s11136-016-1304-8

McClendon, M. J., & O'Brien, D. J. (1988). Question-order effects on the determinants of subjective well-being. *The Public Opinion Quarterly, 52*(3), 351–364. https://doi.org/10.1086/269112

McFarland, S. G. (1981). Effects of question order on survey responses. *The Public Opinion Quarterly, 45*(2), 208–215. https://doi.org/10.1086/268651

Moore, D. W. (2002). Measuring new types of question-order effects: Additive and subtractive. *The Public Opinion Quarterly, 66*(1), 80–91. https://doi.org/10.1086/338631

OECD. (2013). *OECD guidelines on measuring subjective wellbeing*. https://doi.org/10.1787/9789264191655-en

ORC International. (2018). *Gambling prevalence in South Australia (2018)*. https://problemgambling.sa.gov.au/__data/assets/pdf_file/0017/80126/2018-SA-Gambling-Prevalence-Survey-Final-Report-Updated-07.02.19.pdf

Paterson, M., Leslie, P., & Taylor, M. (2019). *2019 ACT Gambling Survey*. https://csrm.cass.anu.edu.au/sites/default/files/docs/2019/10/2019-ACT-Gambling-Survey.pdf

Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist, 54*(2), 93–105. https://doi.org/10.1037/0003-066X.54.2.93

Stark, T. H., Silber, H., Krosnick, J. A., Blom, A. G., Aoyagi, M., Belchior, A., Bosnjak, M., Clement, S. L., John, M., Jónsdóttir, G. A., Lawson, K., Lynn, P., Martinsson, J., Shamshiri-Petersen, D., Tvinnereim, E., & Yu, R.-r. (2018). Generalization of classic question order effects across cultures. *Sociological Methods & Research, 49*(3), 567–602. https://doi.org/10.1177/0049124117747304

Strack, F., Schwarz, N., & Wanke, M. (1991). Semantic and pragmatic aspects of context effects in social and psychological research. *Social Cognition, 9*(1), 111–125. https://doi.org/10.1521/soco.1991.9.1.111

Sudman, S., Bradburn, N. M., & Schwarz, N. (1996). *Thinking about answers: The application of cognitive processes to survey methodology*. Jossey-Bass.

Tourangeau, R., & Rasinski, K. A. (1988). Cognitive processes underlying context effects in attitude measurement. *Psychological Bulletin, 103*(3), 299–314. https://doi.org/10.1037/0033-2909.103.3.299

Tourangeau, R., Singer, E., & Presser, S. (2003). Context effects in attitude surveys: Effects on remote items and impact on predictive validity. *Sociological Methods & Research, 31*(4), 486–513. https://doi.org/10.1177/0049124103251950

****

For correspondence: Kate Sollis, M.Sc, Centre for Social Research and Methods, Australian National University, 146 Ellery Crescent, Acton, ACT, 2601, Australia. E-mail: Kate.Sollis@anu.edu.au