Three Experiments on the Causes of Differences in Estimates of Gambling and Gambling Harm in General Population Surveys

Patrick Sturgis, Department of Methodology, London School of Economics and Political Science

Jouni Kuha, Department of Methodology and Department of Statistics, London School of Economics and Political Science

Shane How, NatCen Social Research

Ioana Maxineaunu, NatCen Social Research

July 2025

Executive Summary

Estimates of gambling and gambling harm vary substantially depending on how surveys are conducted, with online self-completion surveys like the Gambling Survey for Great Britain (GSGB) typically reporting much higher rates than traditional face-to-face surveys such as the Health Survey for England (HSE). These discrepancies have created uncertainty about which mode produces more accurate figures, raising concerns for evidence-based policy and regulatory oversight. This study was commissioned to further understanding of the underlying causes of these differences through controlled experimental testing.

The study tested three hypotheses using a 2x2x2 factorial randomised design with a probability sample drawn from the NatCen panel. It investigated whether: (1) survey invitation wording that explicitly mentions gambling affects who responds, (2) interviewer presence suppresses self-reporting of gambling harm due to social desirability, and (3) the length and specificity of the gambling activities list affects prevalence estimates.

Mentioning gambling explicitly in the survey invitation did not affect the overall response rate but did lead to a 4-percentage point increase in reported gambling participation, suggesting that individuals with a personal interest in gambling were more likely to take part. The PGSI>0 rate was 1.8 points higher in the gambling-invitation group, though this difference did not reach statistical significance. These findings are consistent with longstanding evidence in survey methodology that topic salience influences sample composition in ways that can affect prevalence estimates.

The presence of an interviewer had a substantial effect on reported gambling harm, with PGSI>0 rates 4.4 percentage points higher in the online self-completion condition compared to telephone interviews. This represents an almost 50% increase in reported harm and strongly suggests that respondents under-report undesirable behaviours in the presence of an interviewer. While in-person surveys mitigate this to some extent using self-completion modules, social desirability pressures may still be present due to household dynamics and interviewer presence.

Expanding the list of gambling activities to reflect newer formats, particularly online forms of gambling, led to slightly higher rates of reported gambling but did not significantly change PGSI>0 estimates. The findings align with previous results from the Gambling Commission's own testing, which also found minimal impact from updating the activity list. This suggests that measurement coverage via the activity list does not explain the substantial gaps between older and more recent survey estimates.

The experimental design provides rare causal evidence, isolating the effects of design features from confounding influences. While previous observational studies suggested these factors might play a role, this study offers stronger evidence that both topic framing and survey mode causally influence gambling prevalence estimates. The lack of comparable experimental research in this domain underscores the importance of these findings for methodological best practice.

The study recommends revision of the Gambling Commission's guidance on interpretation of the GSGB's estimates of gambling and gambling harm to better reflect the likely causes of differences between them and those of earlier health surveys given the results of this study. It also recommends conducting detailed benchmarking against the recently published results of the 2023/24 Adult Psychiatric Morbidity Survey.

Introduction

The collection of robust, accurate data on gambling behaviour and harm is central to the Gambling Commission's statutory duty under section 26 of the Gambling Act 2005 to monitor and publish evidence on gambling in Great Britain. Historically, this function was fulfilled through periodic bespoke face-to-face surveys, the British Gambling Prevalence Surveys (BGPS) carried out in 1999, 2007, and 2010. These used random probability sampling and in-person interviews to yield what were then considered best-practice estimates of gambling behaviour and related harms. However, declining response rates and rising fieldwork costs made it increasingly difficult to sustain this approach, particularly in the context of broader public sector spending contraction from 2010 onwards.

In response to these pressures, the Gambling Commission opted to integrate its gambling prevalence monitoring into the Health Surveys for England, Scotland, and Wales. These surveys also use random probability sampling and in-person interviewing, with sensitive modules, such as the Problem Gambling Severity Index (PGSI), administered via self-completion paper questionnaires as part of the interview. The estimates of gambling-related harm derived from these surveys were consistently low. For example, the 2016 Health Survey for England (HSE) found that only 0.7% of the adult population met the PGSI threshold for problem gambling (score 8+), and 4.2% registered a score of 1 or more, indicating some degree of harm.

These figures were at odds with broader concerns at the time around growing gambling harms and anecdotal evidence of rising treatment and service need. In 2019, the gambling charity GambleAware commissioned YouGov to conduct a Treatment and Support Survey using their non-probability online panel. The survey's primary purpose was to explore demand for treatment among those experiencing gambling-related harm, but it necessarily included questions to identify this sub-population via the PGSI. The results showed substantially higher estimates compared to the health surveys, with 2.7% classified as problem gamblers and 13.2% as experiencing harm. These figures were more than three times the HSE estimates from just a year earlier.

The magnitude of this discrepancy prompted further investigation. GambleAware commissioned an independent review (Sturgis 2020) to assess the likely sources of difference between the YouGov and health survey estimates. Drawing on the Total Survey Error framework (Groves and Lyberg 2010), the report considered possible contributions from coverage error, sampling error, measurement differences, and biases due to differential nonresponse. While both surveys had limitations in their designs, the report concluded that the true prevalence of gambling harm likely lay closer to the health survey estimates, largely due to the advantages of probability sampling for valid estimation. However, it also acknowledged that this could not be established definitively, and that the possibility of meaningful underestimation due to measurement error in the face-to-face surveys could not be ruled out.

To investigate the matter more satisfactorily, a follow-up study (Sturgis and Kuha 2021) was commissioned by GambleAware. This compared estimates from eight surveys that fielded the same gambling and PGSI questions within a comparable time window but varied systematically in design features such as sampling frame, sampling design, mode of administration, and questionnaire format. These surveys included the 2016 and 2018 Health Surveys, two successive waves of the YouGov Treatment and Support survey, and specially commissioned surveys conducted by NatCen, Kantar, Ipsos-MORI, and Yonder.

This study found a consistent pattern. Online self-completion surveys, whether based on probability or non-probability samples, produced substantially higher estimates of gambling harm than the health surveys. PGSI+1 estimates (the proportion of the population estimated to score above 0 on the PGSI) ranged from 7% to 16% in the online surveys, compared to just 4% in the health surveys. This pattern of variation enabled the authors to rule out several potential sources of error, including differences in coverage, sampling variability, and questionnaire content.

Sturgis and Kuha (2021) concluded that the most likely explanation for the online to face-to-face divergence in estimates was nonresponse bias. Specifically, individuals experiencing gambling harm appeared to be somewhat under-represented in the face-

to-face surveys, but over-represented in online surveys, particularly those employing panel designs. Both online panel membership and online forms of betting are facilitated by tech-literacy and frequent internet use, and respondents with these characteristics are less likely to attrit from panels.

At the same time, evidence from the health surveys suggested that respondents reported fewer gambling harms when other household members were present during the interview, a pattern that suggested social desirability bias may depress estimates in interviewer-administered surveys. While this pattern did not fully resolve the accuracy question, it suggested that the lower estimates observed in the health surveys may be more downwardly biased than had previously been assumed.

At around this time, the Gambling Commission was reconsidering its approach to delivery of gambling prevalence statistics. The use of the health surveys, while methodologically robust, was increasingly unsatisfactory in terms of responsiveness and control. The Commission lacked flexibility over when gambling modules could be fielded, how questions were prioritised, and how measurement was implemented. In response, and in line with broader trends in survey research, it initiated the development of a bespoke push-to-web survey design, the Gambling Survey for Great Britain (GSGB).

Following extensive consultation and development, the GSGB adopted a methodology that reflects broader developments in social survey practice. It draws a stratified random sample from the Postcode Address File (PAF) and invites up to two adults in each household to complete the questionnaire online, with a paper option available for those unable or unwilling to respond digitally. This push-to-web approach combines the cost-efficiency and privacy of online self-completion with the inferential robustness of probability-based sampling, although response rates for this design are low in comparison to the health (and other in-person) surveys. In an independent review of the GSGB design, commissioned by the Gambling Commission, Professor Sturgis concluded that the development process was exemplary and that the survey represents a high-

quality platform for monitoring gambling behaviour and harm in Great Britain (Sturgis 2024).

As would be expected from the study by Sturgis and Kuha (2021), however, the GSGB produces substantially higher estimates of gambling harm than the health surveys it replaces. This led to a recommendation, in the 2024 review, for further experimental investigation into the methodological drivers of variation in estimates. Specifically, the review identified three priority areas for follow-up: (1) whether the topic framing of survey invitations affects who chooses to participate; (2) the extent to which the questionnaire administration by an interviewer suppresses self-reports of gambling harm due to social desirability bias; and (3) whether the comprehensiveness of the gambling activities list influences the proportion of respondents routed into the PGSI and, therefore, the level of prevalence estimates.

The present report responds to these recommendations. It presents the results of randomised experiments designed to test three hypotheses in a controlled setting. By adopting an experimental design, it is possible to go beyond previous observational comparisons and offer causal evidence about the ways in which survey design features influence estimates of gambling behaviour and harm. We investigate the following questions:

1. Survey Invitation Wording (Topic Salience): Does describing the survey as being about gambling, as opposed to a neutral description such as “health and lifestyle”, affect whether people who gamble chooses to take part, and in turn, the prevalence estimates observed?
2. Mode of Administration (Social Desirability): Does the presence of an interviewer during survey administration lead to lower self-reports of gambling harm, relative to online self-completion?
3. Gambling Activities List (Measurement Coverage): Does a longer, more up-to-date list of gambling activities lead to more respondents being routed into the PGSI module and higher overall estimates of harm?

The experimental design and hypotheses were pre-registered on the Open Science Foundation (OSF) website prior to the start of fieldwork link. Our expectations set out there are the following:

1. Mentioning gambling as the survey topic in the invitation email will yield higher estimates of gambling and gambling harm compared to when the topic is specified as health and lifestyle. It is well known that one of the drivers of survey response is interest in the topic of the survey (Groves, Presser, and Dipko 2004) and it follows from this that people who gamble may be more likely to take part in surveys when the stated topic is gambling compared to other topics. Such a correlation would have the effect of increasing estimates of gambling prevalence in surveys that explicitly mention gambling as the survey topic compared to surveys that do not.

2. Estimates of gambling and gambling harm will be higher in self-completion than in the interviewer-administered mode. Gambling is a normatively undesirable behaviour, particularly harmful gambling, and so some people are likely to find disclosing such behaviour when an interviewer is present embarrassing and will choose to under-report these behaviours. The same effect is likely to occur if an interview is conducted in the presence of other household members (Sturgis and Kuha, 2021). If participants do not disclose their true responses to gambling questions, this would result in surveys which involve interviewers in any capacity[1] under-estimating the prevalence of gambling behaviours, relative to self-completion surveys where no interviewer is present.

3. The new (GSGB) list of gambling activities will yield higher estimates of gambling and gambling harm than the old (HSE) list. The question here is whether estimates of problem gambling, as measured by the PGSI, are higher when the list of questions measuring gambling activity over the previous 12 months includes a broader and more up-to-date set of activities, particularly online gambling (which we will call the new list), compared to the standard set of gambling activities used in the Health Survey for England (HSE) (which we will call the old list). Our expectation is that the new list will

[1] Note that questions on gambling participation and the PGSI are asked in paper or online self-completions in the Health Surveys for England and Scotland, but an interviewer is present at the time of completion.

produce higher estimates of gambling and gambling harm. This is because the old set of questions may miss some participants who gamble online and are therefore not administered the PGSI (which is only administered to respondents who report some gambling in the past year). Because such missed participants may have higher PGSI scores (given the correlation between online gambling and PGSI), the effect would be that the old list produces lower estimates of average PGSI among those who do gamble.

Methodology

To test these hypotheses, we implemented randomised factorial experiments using the NatCen panel, a probability-panel recruited via address-based sampling from the Postcode Address File (PAF). Fieldwork was conducted between 1st May and 28th May 2025. In total 6745 respondents were invited, drawn at random from amongst NatCen panel members who had provided a telephone number. Selection probabilities were adjusted by sampling in proportion to weights reflecting the extent to which panel member characteristics (age, sex, region, household structure, income, education, economic activity, ethnicity, tenure, social class, interest in politics and party support) were over- or under- represented in the eligible panel.

Of these, 3,745 cases were randomly assigned to be interviewed by phone (CATI) and 3000 by online self-completion (or Computer Assisted Web Interview (CAWI)). More cases were issued to the CATI mode because it was anticipated that the response rate would be lower in this condition. This proved to be true, with 1206 interviews completed in the CATI condition, representing a cooperation rate of 32%, while 1,746 interviews were completed in the CAWI condition, representing a cooperation rate of 58%.

All selected participants were invited by email to take part in a survey that would last approximately 15–20 minutes. The invitation text differed by treatment condition. In the control group, the survey was described as being about “Health, wellbeing and recreation” while in the treatment group, it was explicitly described as being about “Betting, lotteries and games.” Full wordings for the survey invitations and reminders can be found in the Appendix.

Those allocated to the online mode were sent a unique survey link via email, clicking on the link took them to the survey landing page. Participants assigned to the telephone mode were sent the invitation email and asked to make contact with NatCen by phone to make an appointment for an interview. Respondents who did not make a call were contacted by a trained interviewer to arrange an interview appointment.

The gambling activity list treatment varied the number and type of gambling activities asked about in the initial screening questions. Half the sample received the standard list used in previous Health Surveys for England. The other half received an extended and updated list reflecting the categories included in the new GSGB instrument.

The treatment assignments for the three different conditions (invitation topic, presence of an interviewer, and list of gambling activities) were independent of each other. As a result, for each of them half of the respondents were assigned to each of the two treatments, and all 2x2x2=8 combinations of different treatments were assigned to around the same number of respondents.

Respondents who reported no gambling activity over the previous 12 months were not administered the PGSI questions, in accordance with standard procedure. All other respondents were administered the full 9-item PGSI. Completing the questionnaire lasted, on average, 10.7 minutes for online respondents and 22.6 minutes for CATI respondents. The full questionnaire can be found in the Appendix.

The results in this report focus on two key outcome variables: 1. whether the respondent reported having gambled in the past year (excluding purchasing National Lottery tickets), and 2. whether the respondent scored above zero on the PGSI. Note that respondents who reported no gambling in the past year are also scored zero on the PGSI, despite not completing it. The rationale for this is that it is conceptually consistent in that they do not experience gambling harms while also maximising sample size.

The univariate estimates for these two variables are broadly in line with those from recent online surveys, with 68% reporting having gambled in the previous 12 months (56% excluding National Lottery tickets) and 10.9% having a PGSI score greater than zero. Note that these estimates are not directly comparable to standard general population surveys because of the restriction to panel members having provided a

telephone number and because the estimates are aggregated over different modes and measurement approaches.

In addition to these experimental outcomes, the online self-completion condition included an observational component. After completing the gambling questions, respondents in the online condition were asked to reflect on whether their answers would have differed had they been interviewed by telephone rather than completing the survey online. This was designed to provide ancillary evidence on the likely direction and magnitude of social desirability effects, though this component is non-experimental and subject to potential demand characteristics.

Analysis and Results

For the invitation topic and gambling list treatments we present the results of crosstabulations of the experimental and outcome variables with Chi-squared tests of association, presented as bar charts. Estimates are weighted to account for differential nonresponse from the NatCen panel and for calibration to population totals. For the mode treatment, there was differential nonresponse, with a higher response rate in the online than the CATI mode. This raises the possibility that differences observed on the outcome variables may be due, in whole or in part, to respondent characteristics that are correlated with response propensity and gambling behaviour. For this experiment, we therefore fit logistic regression models controlling for the following respondent characteristics: age, sex, highest qualification, ethnic group, and frequency of internet use. We then take the mean of the predicted scores from these regressions to estimate the treatment effects. All analyses are conducted in R version 4.5.1.

Experiment 1 Invitation Topic

The first experimental treatment is whether the survey invitation explicitly mentioned gambling or referred to health as the topic of the survey. Figure 1 shows that the invitation content had no significant effect on the probability of responding, with a 41% cooperation rate in each experimental condition. Note that the sample for this analysis is the full issued sample (n=6745).

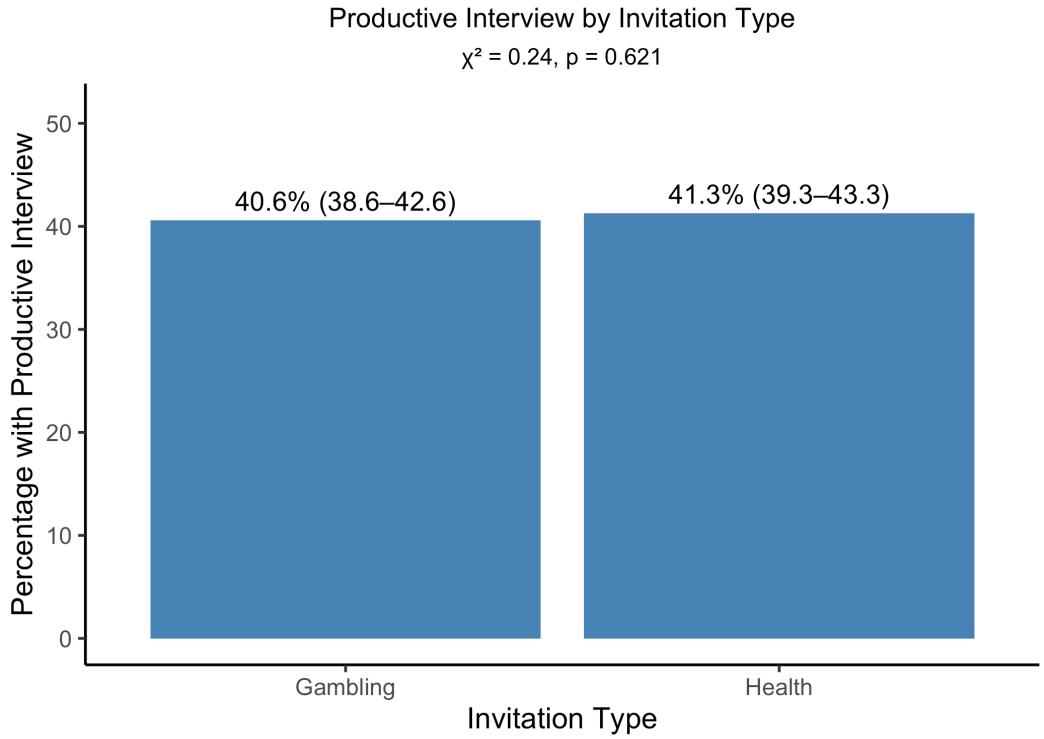

Figure 1 Effect of Invitation Topic on Probability of Responding

It is plausible that some panel members do not read the email invitation closely and, hence, do not receive the treatment. This may particularly be the case for panel members who have completed a large number of surveys and who simply click on the link and complete the survey for the monetary incentive without reading the invitation. We assessed this possibility by fitting the interaction of the invitation treatment with a variable measuring the number of surveys the panel member participated in over the previous year. The interaction was non-significant, providing no support for this expectation.

Turning next to the effect of the specified invitation topic on the level of reported gambling over the past year, Figure 2 shows a significant effect in the expected direction. Panel members who received an invitation explicitly mentioning gambling reported a 4 percentage points higher rate of gambling compared to those whose invitation described the survey as being about health and lifestyle. Because the questionnaires were identical in treatment and control conditions, we conclude that this difference is not a measurement effect but, rather, results from people who gamble being more likely to respond to the survey when the topic is specified as gambling than when it is health and lifestyle. It is possible that priming respondents with different information about the content of the survey caused them to answer the gambling questions differently but this seems unlikely.

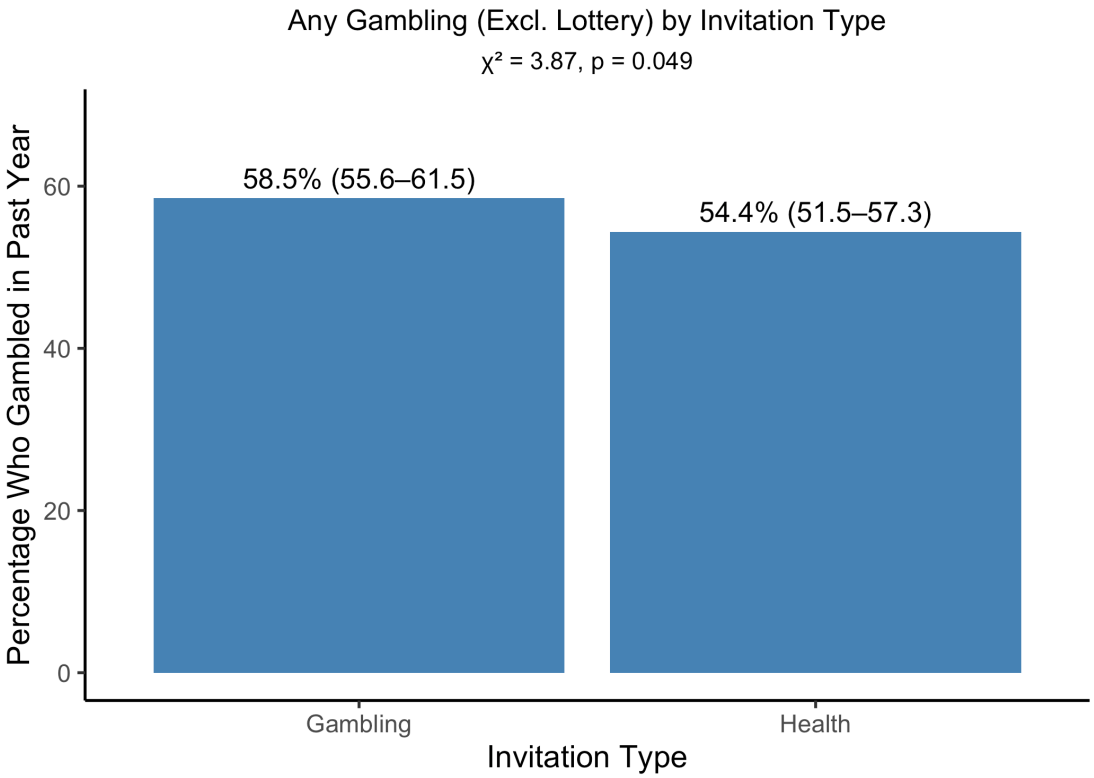

Figure 2 Effect of Invitation Topic on Reported Gambling in the past year

Figure 3 shows the effect of invitation topic of the probability of the respondent scoring 1 or above on the PGSI. While the effect is in the expected direction, with 1.8 percentage points higher in the gambling invitation treatment, the difference is not statistically significant at the $p<0.05$ level, though it should be noted that the power to detect an effect of this size with this sample size is low (0.3).

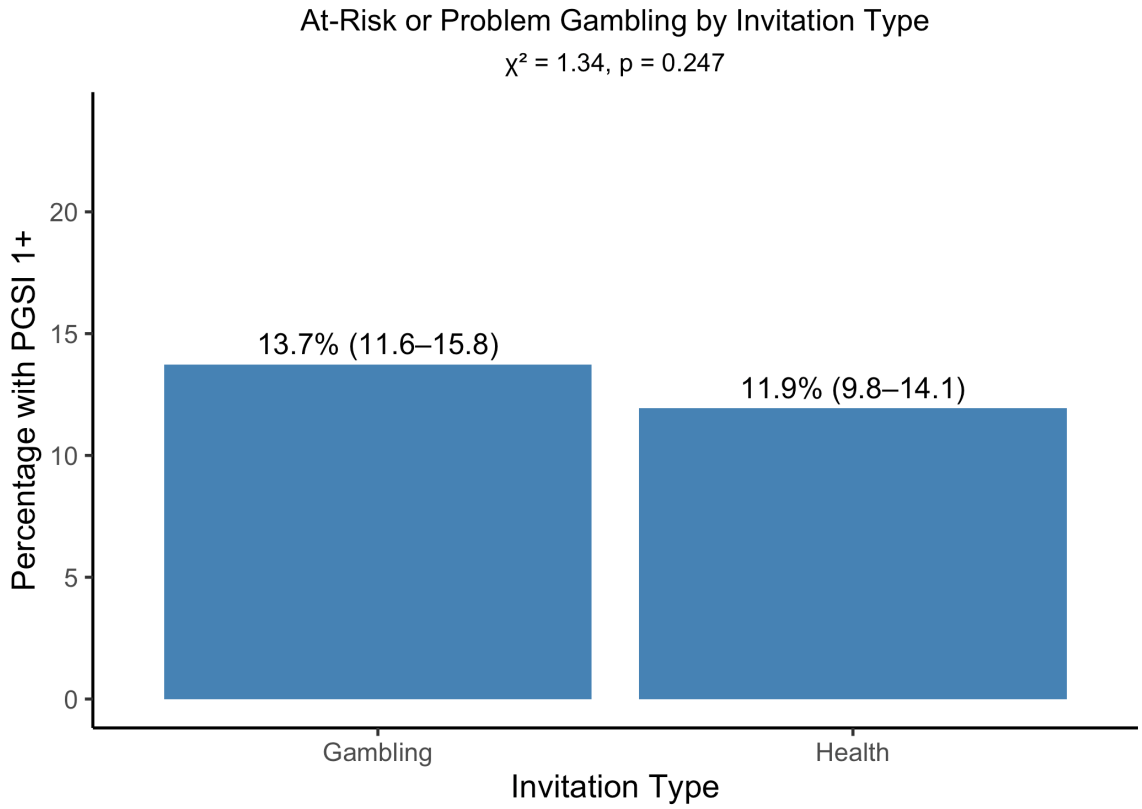

Figure 3 Effect of Invitation Topic on probability of PGSI>0

Experiment 2 – Presence of an Interviewer

Next, we test whether the presence of an interviewer reduced the willingness of respondents to accurately report their level of gambling during the previous year and any harms arising from that. Figure 4 shows that, though the effect is in the expected direction with a 2.8% higher estimate of past year gambling in the online condition, this is marginally outside statistical significance, using a one-tailed test.

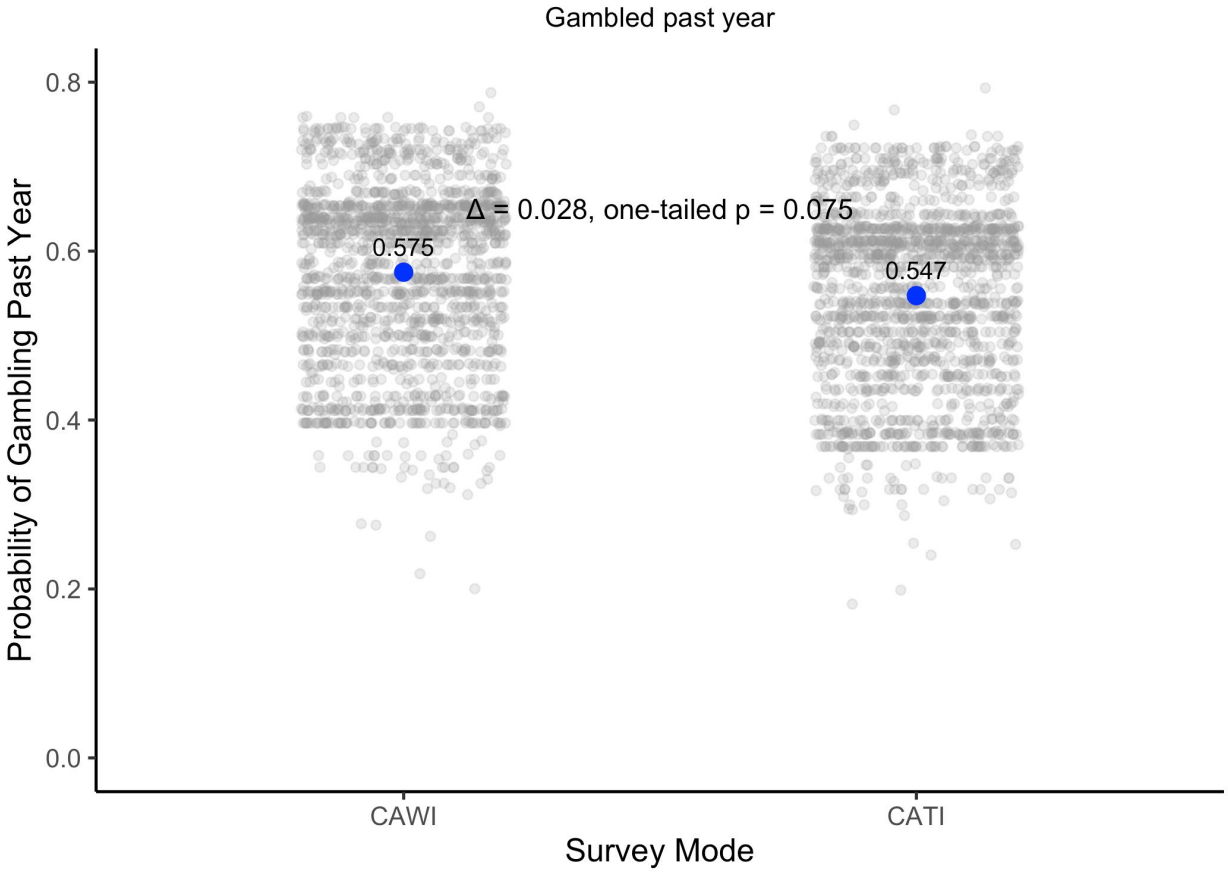

Figure 4 Effect of Interviewer Presence on Probability of Gambling in past year, comparing online self-completion (CAWI) and interviews by phone (CATI). Light grey points are individual fitted values from the model (jittered vertically and horizontally for visibility). Blue dots are the mean of the fitted probabilities.

Figure 5 shows the effect of interviewer presence on the probability of the respondent recording a PGSI score of 1 or more. The size of the effect is large and statistically significant, with a 4.4 percentage point higher estimate in the online condition. This represents close to a 50% reduction in the rate of reporting of gambling harm when the questions are administered by an interviewer compared to online self-completion, a sizeable and substantively important effect.

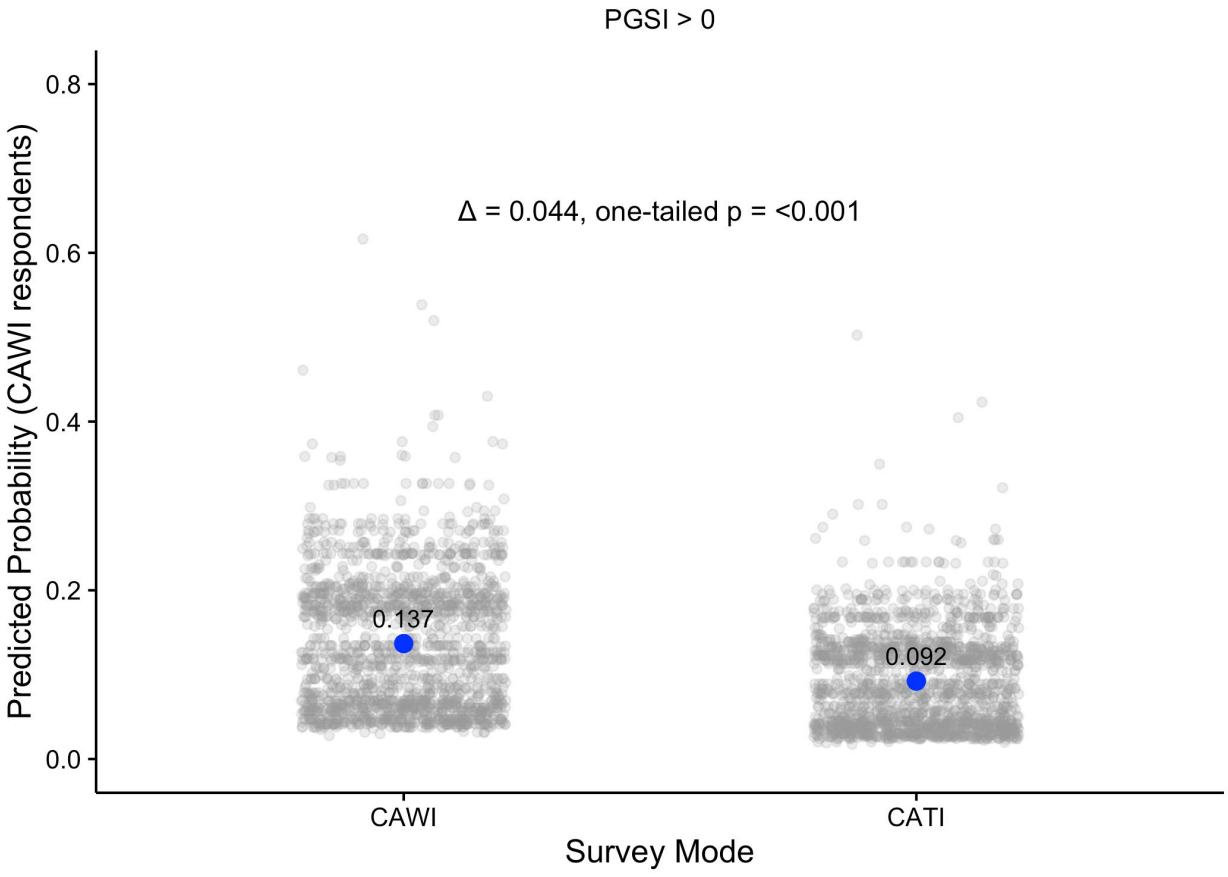

Figure 5 Effect of Interviewer Presence on Probability of PGSI>0, comparing online self-completion (CAWI) and interviews by phone (CATI). Light grey points are individual fitted values from the model (jittered vertically and horizontally for visibility). Blue dots are the mean of the fitted probabilities.

Experiment 3 Old list v New list of Gambling Activities

Lastly, we turn to whether the updated list of gambling activities served to increase the estimated prevalence of gambling and gambling harm. Figure 6 shows an effect in the expected direction for gambling in the past year, with 57.7% reporting having done so with the new list compared to 55.1% with the old list. However, this is not a statistically significant difference. Figure 7 shows only half a percentage point difference between the old and the new list for the PGSI>0 outcome, which is not statistically significant.

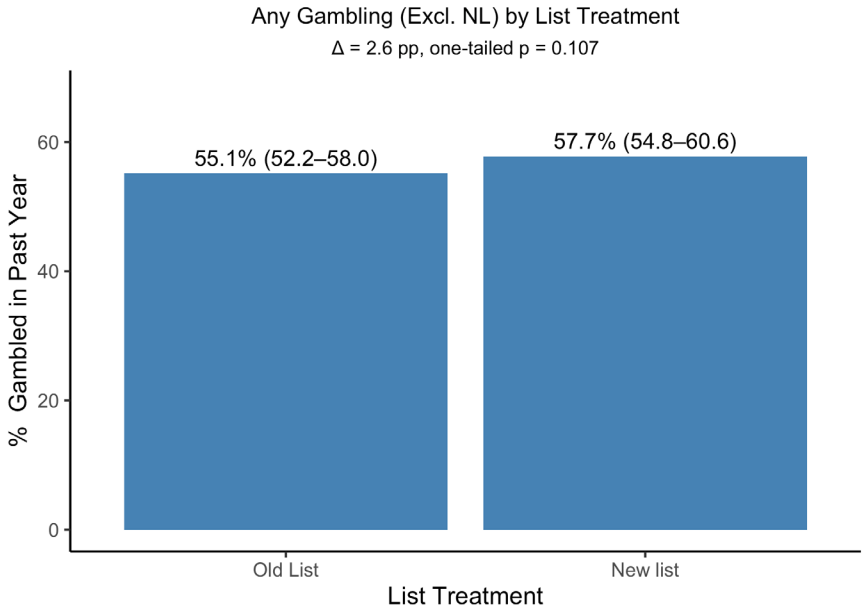

Figure 6 Old v New Gambling Activity List on Gambling Past year

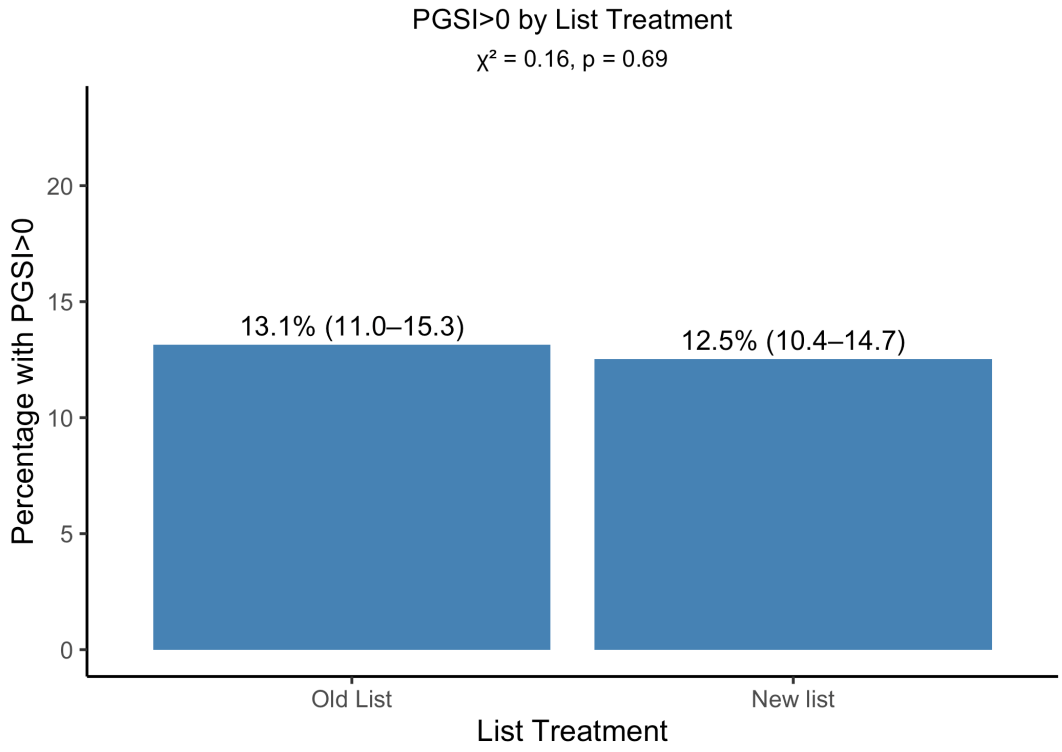

Figure 7 Old v New Gambling Activity List on PGSI>0

After completing the gambling questions, respondents in the self-completion condition were asked to assess whether they would have answered the gambling questions differently had they been asked them by an interviewer. Just 21 respondents (2%) said that they would have answered differently, with the remainder saying they would have given the same answers or that they didn't know how they would have responded. Although this is a small number and, on the face of it, at odds with the experimental results, 18 of the 21 (86%) said they would have reported less gambling with interviewer administration.

Discussion

Over the past two decades social, technological, and economic change has radically transformed the ways in which surveys are conducted. During this period, we have witnessed a shift from the bedrock of general population surveying being conducted by trained interviewers in people's homes, to a position where the vast majority of surveys are implemented through online self-completion. While this has brought benefits, including faster turnarounds, lower per interview costs, and larger sample sizes, it has also been accompanied by lower response rates and substantial breaks in longstanding time series. Where this happens, it is often not clear what the causes of the discontinuity are, nor whether the pre or the post design change estimates are the more accurate.

The GSGB is a case in point. Having until recently met its statutory duty to publish statistics on the prevalence of gambling and associated harms through face-to-face interview surveys, in 2023 the Gambling Commission switched to a push-to-web, mixed-mode, random-probability design. As Professor Sturgis noted in his recent review of the GSGB (Sturgis, 2024), the process through which this transition was implemented was carried out to a very high standard and the survey represents the state of the art in modern household survey design.

That notwithstanding, he also noted that important questions about key gambling estimates remained unresolved, despite several previous investigations of the matter. Most importantly, the reasons why online self-completion surveys tend to obtain substantially higher estimates of gambling and gambling harm than in-person ones were not well understood. Given the importance of these estimates for public policy, this has been an unsatisfactory state of affairs.

The purpose of this research has been to shed new light on this key question through the use of experimental methods. Using a high-quality random probability sample, the study implemented a 2x2x2 factorial design to test the effect of key design features on gambling prevalence estimates. The results go some way to explaining why online surveys such as the GSGB obtain notably higher estimates than in-person surveys such as the Health Survey for England.

First, specifying gambling as the topic of the survey in the invitation to respondents results in significantly higher estimates of gambling harm. It is well-established in survey methodology that people are more likely to take part in surveys when they have some interest in what it is about, with the opposite also being true (cite). The size of the effect for gambling was quite large, at 4%, and, although smaller (2%) and non-significant for PGSI>0, the test on an effect of this magnitude is quite low-powered with the sample size in this study.

The size of this effect is notable given the comparatively 'weak' nature of the treatment, which is to say it is mentioned in an email to panel members used to completing surveys and who may not pay much attention to the specific details of each survey. It would be reasonable to expect a larger effect for a standalone push-to-web survey like the GSGB, where respondents are likely to spend more time reading the more extensive survey materials before deciding whether or not to take part.

While this finding helps to explain differences in estimates between surveys like the GSGB and the HSE, unfortunately, it does not enable a determination of which is the more accurate. This is because we do not know what the true values of gambling and gambling harm are in the population. It may be that by mentioning gambling in the survey invitation, that people who gamble (and experience harm) become *over*-represented in the sample compared to their composition in the population. To better understand how topic interest affects the direction of the bias, it would be necessary to have already observed gambling behaviour on a sample invited to a new survey. This would be a useful approach for future research on this matter.

There were also large and statistically significant effects of interviewer presence on estimates of gambling harm, with the PGSI>0 estimate 50% higher in the self-completion mode compared to a telephone interview. This supports the theoretical expectation, again well-grounded in the survey methodological literature, that social presence reduces the reporting of socially undesirable behaviour (and promotes the reporting of socially desirable behaviour). Of course, some care is required in extrapolating the effect observed here to in-person interview surveys as they are different in important ways from the design implemented here. While the invitation

treatment can be considered quite 'weak', the interviewer presence treatment is possibly somewhat stronger than would be the case for surveys like the Health Survey for England, that use self-completion for sensitive question modules such as gambling.

That being said, there are still reasons to think that a respondent in this situation will still be subject to social desirability pressures; the interviewer will be physically present in the room as will, potentially, other household members, and the respondent may well have concerns about whether the interviewer will see, or in some way be aware of the responses provided through self-completion. For these reasons, while the exact magnitude of the social desirability bias is uncertain, we can be confident in concluding that a good part of the difference that we observe between online and in-person interview surveys in estimates of gambling behaviour is due to downward biases from interviewer administration.

The third experiment produced less notable results. While the longer and more up-to-date list used in the GSGB did produce a slightly higher estimate of having gambled in the past year (2 percentage points), this was not statistically significant and there was no difference in the estimate of PGSI>) between lists. In 2024, the Gambling Commission conducted this same experiment using data from the YouGov panel and this also showed no evidence of higher rates of gambling harm using the updated list. We can therefore conclude with some confidence that the higher estimates of PGSI in more recent surveys is not due to employing the updated list of gambling activities.

In combination, the results of these experiments suggest that a large part of the variation in estimates of gambling and gambling harm between face-to-face interview and online self-completion surveys is due to differences in sample composition resulting from the subject of the invitation and higher levels of socially desirable responding due to interviewer presence. This leads us to recommend that the Gambling Commission should revise its guidance on interpreting the results of the GSGB to better reflect the likely causes of the differences between it and earlier estimates produced by the health surveys.

Mentioning gambling as the topic of the survey, encourages greater levels of participation from people who gamble, a phenomenon that is consistent with theory and

findings from the survey methodological literature. From these data alone, we cannot know whether this effect moves the estimates closer or further away from the true value. For that, it would be necessary to know the PGSI scores for the sample at the point they are invited. However, the characteristics of gamblers are rather overlapping with those typical of survey nonrespondents, in that both groups are more likely to be male, younger, and with lower educational qualifications. Specifying a survey topic that is of interest may therefore serve to increase participation amongst these low propensity groups. On the balance of probabilities, then, it seems reasonable to conclude that specifying gambling as the survey topic improves the accuracy of estimates compared to when the topic is more generally about health and well-being. While the effect sizes we have found here are small at around 2 percentage points, it would be reasonable to expect them to be somewhat larger when the invitations are sent to a fresh push-to-web sample and when the advance materials are more numerous and substantial.

There was a larger effect for social desirability bias, with around a 4 percentage point higher estimate on the PGSI when respondents completed the survey online compared to being interviewed on the phone. While in-person interview surveys mitigate this by using self-completion methods in the survey, this is unlikely to be entirely robust for three reasons. First, a substantial minority of respondents request that the interviewer administer these questions. Second, many respondents my believe that the interviewer will see their responses even though they did not administer them. And, third, other household members will often be present during the interview, lending another – possibly more powerful - social presence to the interview context.

In conclusion, then, it is plausible on the balance of probabilities that self-administration of the PGSI adds 5-6 percentage points to estimates compared to interviewer-administration.

In his 2024 independent review, Professor Sturgis recommended that the Gambling Commission should seek opportunities to benchmark the estimates from the GSGB against a contemporaneous face-to-face interview survey in the future. In June 2025, NHS England published the results of the 2023/4 Adult Psychiatric Morbidity Survey (APMS) which contained the gambling activity and PGSI questions.

The APMS was administered through face-to-face interviewing, with a small (3%) CATI sample. The gambling questions were completed by the respondent during the interview, although 20% requested that they be administered by the interviewer. The response rate for the gambling module was 26%, reflecting the increasing difficulty of obtaining high response rates using this approach. This is not much higher than the GSGB, at 17% in its first year and certainly not a sufficiently large difference to explain the wide divergence in gambling estimates.

The APMS estimated that 4.4% of adults had a PGSI score of 1 or above, compared to 14.3% in the 2023 GSGB. Taking into account the effects observed in this study, we might expect that 5-6 percentage points of this difference can be accounted for by the different survey invitations and modes. However, this still leaves approximately a third of the difference unaccounted for. A second recommendation of this report is, therefore, that the Gambling Commission should undertake further work to benchmark the GSGB against the APMS.

References

Groves, Robert M., and Lars Lyberg. 2010. 'Total Survey Error: Past, Present, and Future'. *Public Opinion Quarterly* 74(5): 849–79. doi:10.1093/poq/nfq065.

Groves, Robert, Stanley Presser, and Susan Dipko. 2004. 'The Role of Topic Interest in Survey Participation Decisions'. *Public Opinion Quarterly* 68(1): 2–31. doi:10.1093/poq/nfh002.

Sturgis, Patrick. 2020. *An Assessment of the Accuracy of Survey Estimates of the Prevalence of Problem Gambling in the United Kingdom*. London: GambleAware.

Sturgis, Patrick. 2024. *Assessment of the Gambling Survey for Great Britain (GSGB)*. London: Gambling Commission.

Sturgis, Patrick, and Jouni Kuha. 2021. *Methodological Factors Affecting Estimates of the Prevalence of Gambling Harm in the United Kingdom: A Multi-Survey Study*. London: GambleAware.

Appendix

TBC