

Measuring Gambling Related Stigma:

A secondary analysis of two validated scales

Full Report

September 2024

Ipsos UK

GambleAware



Contents

1 Executive summary.....3

1.1 Introduction 3

1.2 GESS Overview 3

1.3 GPSS Overview 4

1.4 Opportunities for further research..... 5

2 Introduction and methodology7

2.1 Background..... 7

2.2 Introduction to GPSS and GESS 8

2.3 Application of the scales: Review of literature on GPSS and GESS 9

2.4 Use of GPSS and GESS in stigma reduction campaign evaluation 12

2.5 Methodology for investigation 15

3 Key findings 17

3.1 GESS Key findings 17

3.2 GPSS Key findings 23

4 GESS Technical Annex..... 30

4.1 Assessing structural integrity 30

4.2 Grouping the summary score 31

4.3 Short form development 36

4.4 Regression overview 41

4.5 GESS..... 43

5 GPSS Technical Annex..... 44

5.1 Impact of reductive vs non-reductive language in the questionnaire 44

5.2 Assessing structural integrity 46

5.3 Grouping the summary score 49

5.4 Correspondence between GPSS scales 54

5.5 Short form development 55

5.6 GPSS..... 58

Appendix 59

1 Executive summary

1.1 Introduction

The Gambling Perceived Stigma Scale (GPSS) and the Gambling Experienced Stigma Scale (GESS) provide a tool to better understand the prevalence and experience of gambling-related stigma. Little work has been undertaken to understand how best to interpret the outputs from these scales. Ipsos has been commissioned by GambleAware to conduct a secondary analysis of GPSS and GESS data to help inform the application of these scales as part of GambleAware's wider stigma reduction programme.

Several statistical analysis techniques were used for this investigation, including:

- T-Tests to explore the impact of two different versions of question wording used to capture GPSS.
- Factor Analysis to consider whether items on GPSS and GESS scales contribute equally to the overall measurement.
- Latent Class Analysis to identify different classes (groups) within the total scores used for GPSS and GESS.
- Regression Analysis to determine the relationship between GESS, GPSS and key demographics.

By applying the above statistical analysis techniques to GPSS and GESS survey data, this research has identified:

- How best to classify and interpret the scores of each of the scales;
- Whether the analysis supports the case for a reduced length (short form) questionnaire for either of the scales;
- The optimum design for a short form questionnaire for the scales.

Further analysis has identified strength of association between demographic characteristics and scores on each of the scales.

1.2 GESS Overview

The Gambling Experienced Stigma Scale (GESS) measures the self-perceived gambling-related stigma experienced by those who gamble.¹ The secondary analysis conducted for this study found that:

- **GESS measures a single underlying dimension of experienced stigma, and thus the 'total' score is an appropriate tool for tracking and evaluation.** All items measured by the scale

¹ Andrà, C., Priolo, G., Merlin, F. *et al.* Differences in Perceived and Experienced Stigma Between Problematic Gamblers and Non-gamblers in a General Population Survey. *J Gambli Stud* 38, 333–351 (2022).

contribute equally to the score; thus, a lower score reliably indicates lower experience of stigma, and a higher score reliably indicates higher experience.

- **There are four natural groupings of experience of stigma as measured by GESS: low (GESS score of 13-19), moderate (score of 20-29), high (score of 30-39) and very high (score of 40+).** There is a large floor effect, where a GESS score of '13' accounts for a large proportion of participants ('strongly disagree' to all statement). Though there is merit in separately monitoring change over time of the score '13', a five-class grouping would be less statistically robust than the four-class grouping and is therefore less suitable for developing a short form of the measure.
- **A short form of GESS is statistically viable; 3-, 4- and 5-item short forms of GESS are all appropriate monitoring tools, but trade off accuracy and respondent burden.** The 5-item scale provides the highest degree of accuracy and allows for tracking of items relevant to the Stigma Campaign.
- **Socio-demographics are a strong predictor of GESS scores.** Experience of stigma is greatest among young religious males. This group make up 8.1% of those who gamble and have an average GESS scores of 35.5 (compared to an average of 27.9).
- **There is a close relationship between socio-demographic predictors of GESS and socio-demographic predictors of PGSI.** When controlling for PGSI within a regression analysis, other socio-demographics account for just 1.3% of further variation in GESS scores. This suggests that the same socio-demographic variables predict both PGSI and GESS.

1.3 GPSS Overview

The Gambling Perceived Severity Scale (GPSS) measures the perception of gambling-related stigma at a societal level by the general population.² Key findings from secondary analysis showed that:

- **Future application of the GPSS should continue to use non-reductive language.** While replacing reductive with non-reductive language in the questionnaire had a minimal impact on measuring perceived stigma, the adoption of non-reductive language is important in addressing stigma more generally. As such, it is appropriate to adapt the question wording using non-reductive language.
- **GambleAware should continue to monitor trends in Ostracism and Contempt independently of each other,** rather than monitor change in the GPSS score overall. A substantial minority of people have slightly mixed perceptions towards Ostracism and Contempt, with poor correlation between the two subscales.
- **A three-class model for each subscale within GPSS is optimal.** Latent Class Analysis shows that classifying total scores into smaller groups is a valid exercise; a three-class model performed better, with the four-class model being unstable. Ostracism can be grouped into Low (score of 6-

² ibid

15), Moderate (score of 16-20) and High (score of 21-24); Contempt can be grouped into Low (score of 7-16), Moderate (score of 17-21), and High (score of 22-28).

- **Results of the analysis do not support the proposal of a short-form for the GPSS.** The creation of a short-form GPSS trades off respondent burden with quality and accuracy. Short-form versions of the Contempt and Ostracism scales are not advised due to a significant level of misclassifications between different classifications.
- **Socio-demographics are not a strong predictor of perceived stigma.** Although some socio-economic groups are more likely to show a higher perceived gambling stigma than others, regression analysis shows that socio-demographic characteristics are, at best, only weakly related to Contempt and Ostracism scores.

1.4 Opportunities for further research

- **Further validation against other data or modes is welcome.** Analysis was conducted on a limited sample that may lack generalisability. The data analysis was based on a sample of 4,005 adults in GB within an online survey, so may not be replicable to other jurisdictions, samples and/or forms of data collection. Further research is needed within other datasets and contexts to validate the recommendations outlined within the report. In general, there was a lack of research using these scales, thus more investigation would be useful to add to the evidence base.
- **It will be important to monitor changes in stigma measurements in the context of impact of stigma reduction activities.** It is unclear how measures may change over time. The data used for this study showed a small change from the first wave of data to the second wave. Longer term, it is unclear how sensitive GPSS and GESS will be in monitoring wider shifts in society. As such, it is important to continue to monitor these measures over time to see if stigma reduction activities can effectively impact such wide measures of stigma influenced by external factors (e.g., wider policy and regulation, marketing by gambling companies).
- **Further research should explore comparisons between perceived and experienced stigma.** Although the GESS was strongly linked to demographics, this was far less the case between the GPSS and demographics. It might be that individuals are projecting what they think is the norm (rather than their own views) which makes scores consistent across different groups within society. However, there was some evidence that certain groups who experience the most stigma have internalised their stigma and project it to others within their community (e.g., those experiencing 'problem gambling'). Further research on the role of internalised stigma on perceived stigma may be useful to better understand how to reduce gambling stigma.
- **Further research could also explore whether there is a potential ordering affect between the two scales,** and the extent to which showing GPSS prior to GESS has an impact on the responses given.
- **Further research should compare the composition and structure of GESS and GPSS by different demographic groups.** Multi-group factor analysis could test the validity of the scale by gender, cultural background, age and PGSI status.
- **Further research should explore the role of personal attitudes, values and belief in perceived stigma.** Although there are correlations between demographic groups and

GPSS/GESS scores, analysis suggests these may be explained by other factors, such as personal attitudes and values. Exploration of this could inform future campaign messages and treatment programmes.

2 Introduction and methodology

2.1 Background

A core GambleAware ambition is to reduce the stigma surrounding gambling harms. Stigma is important to counter as it prevents those experiencing harm from seeking support and treatment whilst exacerbating harm in itself (e.g., by driving feelings of guilt/shame, lower feelings of self-esteem and reduce self-efficacy).³ Ipsos research (2023) found that whilst around a fifth of adults (21%) think they know someone who has experienced gambling harms, most are reluctant to speak to them about it for reasons relating to stigma and discrimination (61%). This includes concerns around further stigmatising the person experiencing harms but also attitudes towards those who gamble.⁴ GambleAware ran its first burst of the national “Let’s open up about gambling” public health communication campaign from 13th April to 11th June 2023 to address the effects of this stigma.⁵

The campaign evaluation included two validated scales which seek to measure stigma: the Gambling Perceived Stigma Scale (GPSS) and the Gambling Experienced Stigma Scale (GESS).⁶ These measure the perception of gambling stigma within society, and the extent to which those who gamble feel stigmatised. The GPSS and GESS were developed in 2015 by Donaldson et al to measure perceived (GPSS) and experienced (GESS) stigma related to gambling to provide researchers and others with a tool to measure gambling-related stigma to address the impact of stigma on conducting research into gambling including treatment seeking.⁷

The GPSS and GESS each have 13 items, with each item scored using a four-point Likert scale (ranging from “strongly disagree” to “strongly agree”), the sum of which can be used to create a total score. However, little work has been undertaken to understand how best to interpret these figures as a measure of stigma beyond its original intention as a control for its effect on other measures of interest.

Ipsos has been commissioned by GambleAware to conduct additional secondary analysis on GPSS and GESS data to help inform the future application and utilisation of these scales as part of GambleAware’s wider stigma reduction programme.⁸ The aims of this project were to:

- Test the relationship between individual items and overall scale scores to confirm the construct validity of the GESS and GPSS scales.

³ Quigley, L. (2022) Gambling Disorder and Stigma: Opportunities for Treatment and Prevention. *Curr Addict Rep*, 9(4), 410–419.

⁴ Ipsos (2023) GambleAware Stigma Polling: Key findings. Accessed June 2024. Available at:

<https://www.ipsos.com/sites/default/files/ct/news/documents/2023-04/Gambling%20Stigma%20Polling%20Report.pdf>

⁵ Further information on the background and context to the Stigma Campaign can be found here: Campaign to reduce the stigma associated with gambling harm: summary of background research and data insights. /<https://www.begambleaware.org/sites/default/files/2023-04/Summary%20of%20background%20research%20and%20data%20insights%20for%20stigma%20reduction%20campaign.pdf>

⁶ GambleAware (2023). ‘Let’s Open Up About Gambling’ stigma reduction campaign evaluation. Accessed June 2024.

Available at: <https://www.begambleaware.org/sites/default/files/2023-10/Stigma%20reduction%20campaign%20evaluation.pdf>

⁷ Donaldson, P., Langham, E., Best, T., & Browne, M. (2015). Validation of the Gambling Perceived Stigma Scale (GPSS) and the Gambling Experienced Stigma Scale (GESS). *Journal of Gambling Issues*, 31, 162–199.

⁸ This includes the analysis of data resulting from the use of two variations of the GPSS questionnaire; one using reductive language (e.g., “problem gambler”) and the other, non-reductive language (e.g., “those experiencing harms from gambling”). The two sets of data were compared to understand the impact (if any) of varying the language, with the aim of combining the two data sets together into one combined GPSS dataset. This is discussed in section 1.3.

- Identify the most appropriate way for interpreting the total GPSS and GESS scores; including the creation of meaningful groups/segments that can be tracked over time.
- Develop a short form for GESS and GPSS (i.e., a questionnaire with fewer than the 13 items for each scale) that replicates as accurately as possible the degree of stigma measured by each of the full 13-item forms.
- Next, create meaningful groups/segments suitable for total shortform scores and assess the extent to which the short form accurately predict the correct grouping/segment for each score.
- Explore the associations between sociodemographic characteristics and summary scale scores.

2.2 Introduction to GPSS and GESS

The GPSS and GESS were initially created to provide a better understanding of the impact of gambling-related stigma on treatment seeking. A large body of research had demonstrated that stigma was a barrier to seeking treatment and that it interfered with accurately measuring the prevalence of ‘problem gambling’ (based on a score of 8+ on the Problem Gambling Severity Index Scale).⁹ The researchers aimed therefore to develop scales to measure stigma related to gambling behaviour.¹⁰ The scales were therefore originally conceived and designed as explanatory or control variables to test the impact on other gambling-related outcomes.

Having reviewed existing measures of other stigmatised conditions that shared characteristics with gambling (such as alcohol misuse and eating disorders based on a theoretical framework of stigma). Donaldson et al (2015) identified and adapted items from those measures, tailoring them to gambling. These were developed into a study of 1,370 Australian adults, of which a split sample (n682 and 688) were shown alternative forms of the 18 GPSS items, targeted towards either “recreational gamblers” or “problem gamblers”. Those who said they had gambled within the last 12 months (excluding lottery, scratch cards or raffles) were then shown 18 items of the GESS. This resulted in the creation of two survey instruments, the GPSS and GESS.

The resulting GPSS comprises two dimensions (“Contempt”, measured by 7 items and “Ostracism” measured by 6 items) of perceived stigma while the GESS comprises one dimension of experienced stigma. Each scale measured responses (4-point Likert scale from “strongly disagree” to “strongly agree”) to 13 items.

The original GPSS questionnaire used reductive language (i.e. “problem gambler”) in one of the versions tested (the other version was “gambler”) with a split sample to determine whether language had an impact on responses. However, to better align with GambleAware’s aim of reducing stigmatising language in society¹¹, non-reductive (i.e. “people experiencing gambling harms”) language was tested, the results of which are considered in this report.

⁹ Ferris, J. and Wynne, H (2001). The Canadian Problem Gambling Index: Final Report. Accessed June 2024. Available at: [https://www.greco.ca/Modules/EvidenceCentre/files/Ferris%20et%20al\(2001\)The_Canadian_Problem_Gambling_Index.pdf](https://www.greco.ca/Modules/EvidenceCentre/files/Ferris%20et%20al(2001)The_Canadian_Problem_Gambling_Index.pdf)

¹⁰ Donaldson, P., Langham, E., Best, T., & Browne, M. (2015). Validation of the Gambling Perceived Stigma Scale (GPSS) and the Gambling Experienced Stigma Scale (GESS). *Journal of Gambling Issues*, 31, 162–199.

¹¹ GambleAware (2023). 12 ways to reduce stigma when discussing gambling harms – a language guide. Accessed June 2024. Available at: <https://www.begambleaware.org/sites/default/files/2023-04/Stigma%20Language%20guide.pdf>

2.3 Application of the scales: Review of literature on GPSS and GESS

To support secondary analysis of the GPSS and GESS, a review of the available literature was carried out to find out about the use of the scales outside of GambleAware. Specifically, the aims were to:

- Identify whether the scales have been validated by others.
- Understand how the scales have been used in research.
- Identify whether this research, or any other activities, have resulted in creating any classifications based on participants' scores.

Search terms were: GPSS gambling, GESS, Gambling Perceived Stigma Scale, Gambling Experience Stigma Scale, gambling stigma, GPSS validation, GPSS categories.

Validation of the scales

The review found one piece of research, "[Developing and validating a scale to measure the enacted and felt stigma of gambling](#)" by Donaldson et al (2015) on validating the GPSS and GESS scales. The research aim was to develop a scale to measure stigma related to gambling behaviour that could be used to support the understanding of gambling experiences and the impact of gambling.¹²

The findings supported a single dimensional model of experienced stigma and two-dimensional (namely "Contempt" and "Ostracism") model of perceived stigma. The scales were found to be reliable and robust. In addition, the findings showed that the experienced and perceived stigma scales reliably distinguished between recreational and problem gambling to understand the extent of gambling behaviour and its impact on stigma. However, the research did not explore the sequencing of the scales within the questionnaire and the possible implications of rotating the ordering in which the scales are shown on responses.

Use of GPSS/GESS in research

The scales have been used in research about gambling stigma (or gambling stigma among other gambling harms). Most recently Andra et al (2022) used GPSS and GESS (along with three other gambling harms tools) to test four hypotheses relating to perceived and/or experienced stigma.^{13 14} These were: the relationship between perceived stigma and socio-demographic variables, measurements of perceived stigma between those who gamble and non-gamblers, the relationship (among those who gamble) between high levels of perceived stigma and an increase in experienced stigma and the effect of this on seeking help, and the correlation between experienced stigma and the severity of gambling behaviour.

¹² Donaldson, P., Langham, E., Best, T., & Browne, M. (2015). Validation of the Gambling Perceived Stigma Scale (GPSS) and the Gambling Experienced Stigma Scale (GESS). *Journal of Gambling Issues*, 31, 162–199

¹³ Andrà, C., Priolo, G., Merlin, F. (2022) Differences in Perceived and Experienced Stigma Between Problematic Gamblers and Non-gamblers in a General Population Survey. *J Gambli Stud* 38, 333–351

¹⁴ Andrea et al (2022) used the same sample composition as in Donaldson et al's validation of the GPSS and GESS scale in 2015.

Analysis of GPSS scores included comparing mean scores and the discussion of differences between the research target populations based on particular questions in the GPSS/GESS questionnaires but did not result in any groupings, categories, types etc.

The research found that gambling stigma endured across different socio-demographic groups and that there was a correlation between higher levels of self-perceived stigma, “more severe disorder” and lower likelihood of seeking help.

The report recommended considering education as a route to reducing social stigma towards gambling, with the aim also of promoting treatment as more socially acceptable. The authors were interested in the potential for future research on how social stigma varies between different target populations such as, “non-problematic gamblers”, “non-gamblers”, those who know “problem gamblers” and those who do not. They were also interested in extending the GESS and GPSS to better understand social stigma in the context of the relationship between those who stigmatise others and those who are stigmatised.

Using the GPSS and GESS scales, Leslie (2022) examined relationships between preferred gambling mode, gambling-related stigma, and help-seeking behaviours in a sample of “disordered gamblers”. The GPSS and GESS questionnaires were administered to participants (n517 across three gambling modes) to produce a total score for Contempt and Ostracism and experience of stigma, with higher scores indicating greater endorsement. Among the findings, the research showed experience of stigma was a positive predictor of help-seeking, regardless of preferred gambling mode.¹⁵ Leslie (2022) also researched whether the experience and perception of stigma differ based on preferred gambling mode, using the GPSS and GESS scales.¹⁶ The research found no differences between groups in terms of their perception, or experience (i.e., internalisation) of gambling-related stigma (though there were differences in other gambling harms).

Bailey (2018) explored whether stigma surrounding behaviours related to impulse control disorders (ICD) (e.g., gambling, substance use, eating) resulting from the use of dopaminergic therapies used to treat Parkinson’s disease would influence reporting of presence and severity of ICD to healthcare providers.¹⁷ The sample of 13 Parkinson’s disease patients completed the GPSS scale in addition to the Questionnaire for Impulsive Compulsive Disorders Rate Scale (QUIP-RS) and questionnaires for other compulsive conditions and behaviours to understand the association between conditions and behaviours common among Parkinson’s disease patients with ICD, self and perceived stigma and treatment. The research found a positive correlation between the use of dopamine treatments and the QUIP-RS score but no significant relationship between scores for the GPSS (or measures of other conditions and behaviours) and the QUIP-RS. The author attributed this to the small sample size.

Delfabbro et al (2022) examined the relationship between perceived stigma (GPSS) towards “problem gamblers” and gambling attitudes and other factors, using a sample of 1,787 university students aged

¹⁵ Leslie, R. D (2022) An investigation of the relationship between gambling-related stigma and help-seeking behaviours among online, offline, and mixed-mode disordered gamblers (Master’s thesis, University of Calgary, Calgary, Canada). Accessed July 2024. Available at: <https://prism.ucalgary.ca/server/api/core/bitstreams/0313c627-e3c7-42d1-9097-d27cd603d1d6/content>

¹⁶ Leslie, R. D (2022). A profile of disordered gamblers based on their preferred gambling mode – preliminary results. Accessed July 2024. Available at: https://www.ucalgary.ca/sites/default/files/teams/320/Leslie_AGRI_2022_Poster.pdf

¹⁷ Bailey, M (2018). Barriers to patient report of impulse control disorders in Parkinson’s disease. Accessed July 2024. Available at: <https://www.proquest.com/openview/55ac63ede7d6b832e81ff8be6b61c4b2/1?pq-origsite=gscholar&cbl=18750>

18-30 across the four countries mentioned.¹⁸ The research found some variance in perceived stigma (GPSS) scores, associated in particular with people’s perceptions of the social accessibility of gambling and their broader attitudes to gambling. It found PGSI to have a lesser impact on perceived stigma.

Summary

Although research by Donaldson et al (2015) concluded that the GPSS and GESS are robust, reliable scales with potential for wider use in understanding gambling stigma and barriers to treatment seeking, the scales have had limited use in research since their inception.¹⁹ The scales have aided academic research related to the stigma surrounded gambling harms and reinforced (and to a some extent, slightly expanded) understanding the experience and attitudes to people who gamble. However, no further research has been conducted to explore the utility of the GPSS and GESS in broader application beyond the original intent as a measurement tool for gambling-related stigma, such as being used as an outcome measure. There have not been attempts to define high and low scores and determine the appropriateness and purpose of creating groupings, or explore how these interact with other factors such as demographics (e.g., age, socio-economic group etc) or behaviours.

Table 2.1: GPSS and GESS questionnaire statements

<p>GPSS question: We are interested in your thoughts about people who gamble. For each of the following statements, please consider how you think people who gamble are generally perceived by others.²⁰</p>
<p>Coding: Strongly Disagree (1) Somewhat Disagree (2) Somewhat Agree (3) Strongly Agree (4)</p>
<p>Contempt Subscale</p> <ol style="list-style-type: none"> 1. Most people think [problem gamblers / people experiencing harms from gambling] are liars 2. Once they know a person is a [problem gambler / experiencing harms from gambling], most people will take his or her opinion less seriously 3. Most people think that [problem gamblers / people experiencing harms from gambling] tend to be unreliable 4. Most people think that [problem gamblers / people experiencing harms from gambling] are unable to handle responsibility 5. Most people think that [problem gamblers / people experiencing harms from gambling] are lazy 6. Most people think that problem gamblers / people experiencing harms from gambling] are greedy 7. Most people believe that [problem gamblers / people experiencing harms from gambling] have no self-control <p>Ostracism Subscale</p> <ol style="list-style-type: none"> 8. Many people would be uncomfortable communicating with a [problem gambler / people experiencing harms from gambling] 9. Most people think less of a [problem gambler / people experiencing harms from gambling] 10. Most people would not hire a [problem gambler / people experiencing harms from gambling] to take care of their children

¹⁸ Delfabbro, P., Hundric, D.D., Ricijas, N. (2022) What Contributes to Public Stigma Towards Problem Gambling?: A Comparative Analysis of University Students in Australia, Canada, Croatia and Israel. J Gambl Stud 38, 1127–1141

¹⁹ Donaldson, P., Langham, E., Best, T., & Browne, M. (2015). Validation of the Gambling Perceived Stigma Scale (GPSS) and the Gambling Experienced Stigma Scale (GESS). Journal of Gambling Issues, 31, 162–199

²⁰ As part of a wider commitment to removing use of stigmatising language, GambleAware included a text edit within the GPSS scale. For half the sample, ‘problem gambler’ was replaced with ‘person experiencing harms from gambling’. The original GPSS scale published by Donaldson et al (2015) inserted ‘recreational gambling’ and ‘problem gambling’ substitutions instead.

11. Most people would be suspicious of a person if they knew they were a [problem gambler / people experiencing harms from gambling]
12. Most people would not want to enter into a committed relationship with someone they knew [had a gambling problem / was experiencing harms from gambling]
13. Many people would avoid a person who [had a gambling problem / was experiencing harms from gambling]

GESS question: We are interested in your thoughts about your own gambling experiences. Please indicate how much you agree with each of the following statements.

Coding: Strongly Disagree (1) Somewhat Disagree (2) Somewhat Agree (3) Strongly Agree (4)

1. I feel the need to hide my gambling from my friends
2. I sometimes have the thought that I've screwed up my life by gambling
3. Most people would always suspect that I'd returned to gambling, even if I didn't gamble anymore
4. People have insulted me because of my gambling
5. I have the thought that I should be ashamed of myself for my gambling
6. People can tell that I am a gambler by the way I look
7. Others think I am not worth the investment of time and resources because I am a gambler
8. I sometimes have the thought that I deserve the bad things that have happened to me in life because I gamble
9. I feel the stress in my life is what causes me to gamble
10. Others view me as morally weak because I am a gambler
11. I avoid situations where another person might have to depend on me
12. I don't think I can be trusted because I gamble
13. Once they know I'm a gambler, most people will take my opinion less seriously

2.4 Use of GPSS and GESS in stigma reduction campaign evaluation

GambleAware commissioned Ipsos to conduct an evaluation of its "Let's open up about gambling" stigma reduction above the line media campaign which was live between 13 April to 11 June 2023. Baseline data (wave 1) was collected using an online survey between 29 March and 11 April 2023; a second survey wave (wave 2) was conducted from 9 to 22 June 2023.

Fieldwork was carried out using the Ipsos Online Access Panel (therefore, only reflecting people with digital access). Data was weighted back to age, gender, work status and region (based on the population of Great Britain) to maximise consistency across the waves.

The total sample used for the analysis of GPSS and GESS comprised of:

- A general population sample of 4,005, across 2 waves of research, asked the GPSS.
- A sub-sample of 2,559 aged 18-44 who have gambled in the past four weeks, asked the GESS.

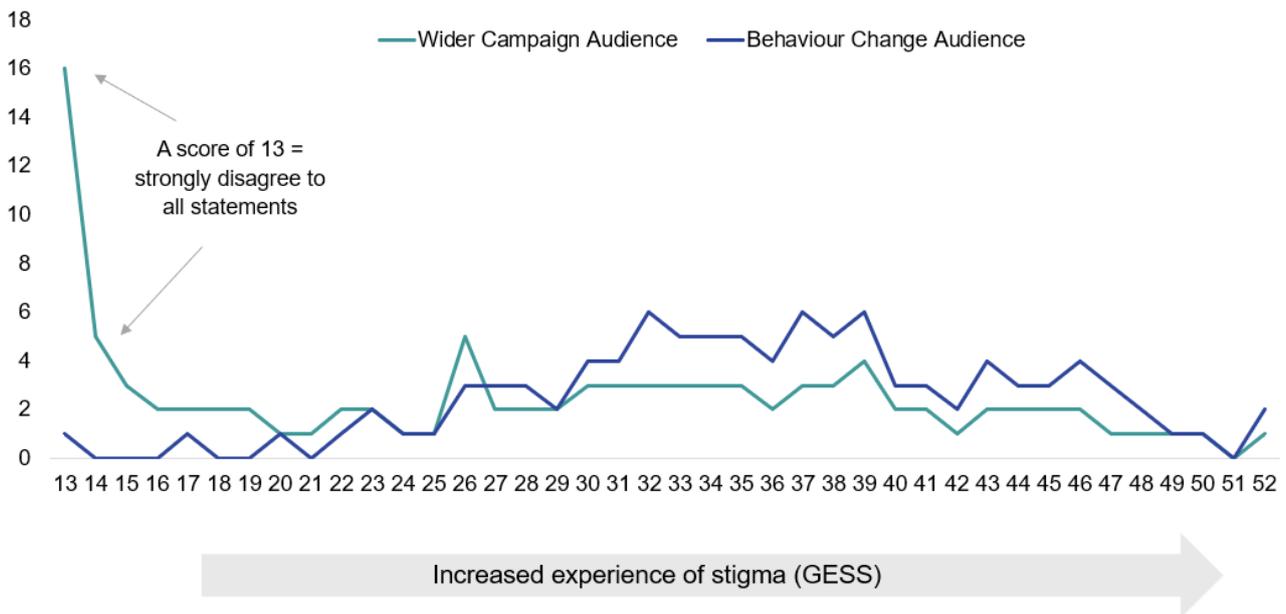
The GPSS and GESS measures were used as part of a wider evaluation framework which monitored the impact of the campaign. The scales provide a robust measure for assessing how well the campaign challenged or overcame gambling stigma among the key and overhear audiences and among specific demographic groups. It is recognised that both perceived and experienced stigma are complex realities that will take time to shift. As such, it should be noted that GambleAware did not expect to see significant change after the first initial bursts of the campaign. Nonetheless, given the multi-year long term

campaign strategy, GESS and GPSS provided a useful set of indicators for monitoring long-term outcomes.

Distribution of scores on the GESS

The GESS scale was asked only of those aged 18-44 who have gambled in the past four weeks (referred to as the “Wider Campaign Audience”). The average mean GESS total score among those aged 18-44 who have gambled in the past four weeks in the baseline survey (wave 1) was 27.9, rising to 29.9 in the post-campaign survey (wave 2). Experience of stigma was higher among those who are experiencing at least moderate problems from gambling (referred to as the “Behaviour Change Audience” and defined as those scoring PGSI 3+ on the PGSI scale), with a mean score of 35.9 and 36.6 between waves. The graph below depicts the distribution from the baseline survey, which shows that a large proportion of participants scored 13 (“strongly disagree”) to all statements. The analysis considered how best to take this into account when developing the groupings of scores for classification of levels of stigma.

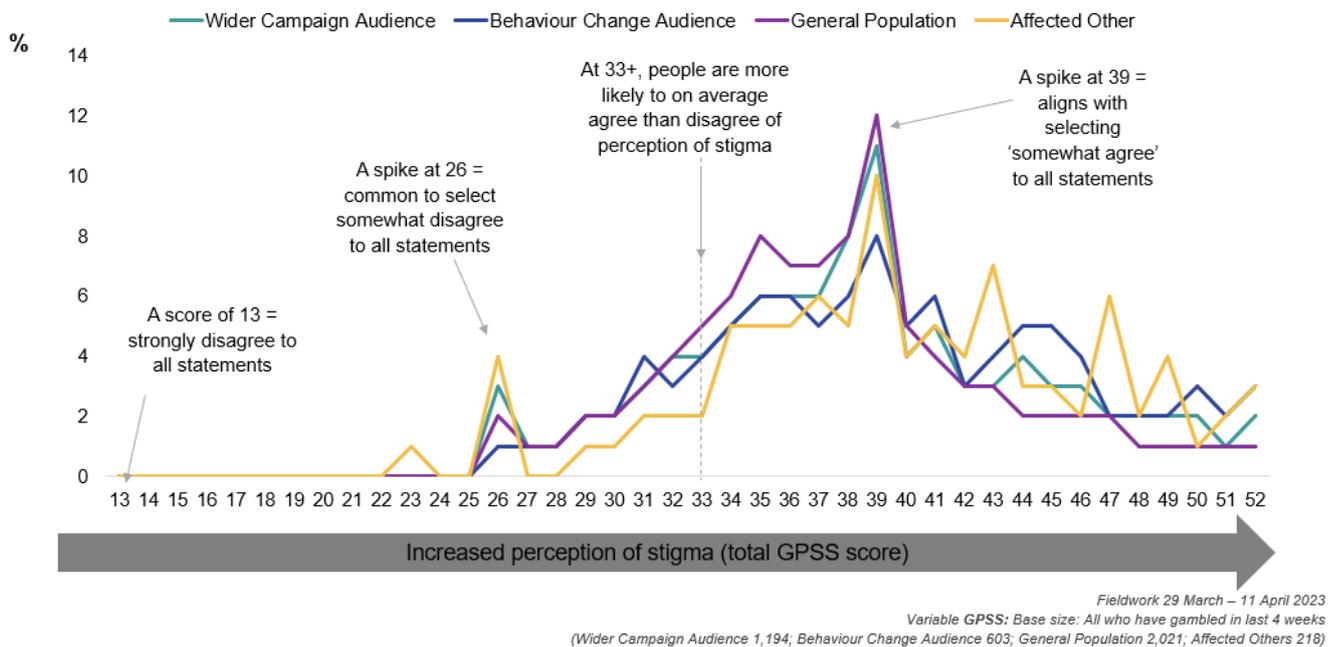
Figure 2.1: Experience of stigma by campaign audience



Source: Ipsos/GambleAware 2023

Distribution of scores for GPSS

The GPSS questions was asked of everyone (all adults aged 18+). The mean GPSS total scores in wave 1 remained largely the same across all key audience groups: 36.4 among the general population, and 37.1 among the “Wider Campaign Audience”. Mean scores in wave 2 were broadly in line (37.6 and 39.4 respectively). The graph below shows that the distribution of GPSS scores spikes at selecting “somewhat agree” to all statements.

Figure 2.2: Perception of stigma, by campaign audience

Experimental design to investigate impact of reductive vs non-reductive language

The GPSS questionnaire experimented with varying the language used in the questionnaire statements. Half the sample was asked a set of questions using “reductive” (“problem gambler”) language while the other half received the same set of questions using “non-reductive” (“experiencing harms from gambling”) language.²¹ It is hoped that there can be a transition to the preferred non-reductive language in future.²² This will add value to understanding the experience of people who gamble and to provide a more general societal perspective on the stigmatisation of gambling. As the impact of this experimental design had implications for how to proceed, analysis was carried out to compare the resulting data from the two subsamples.

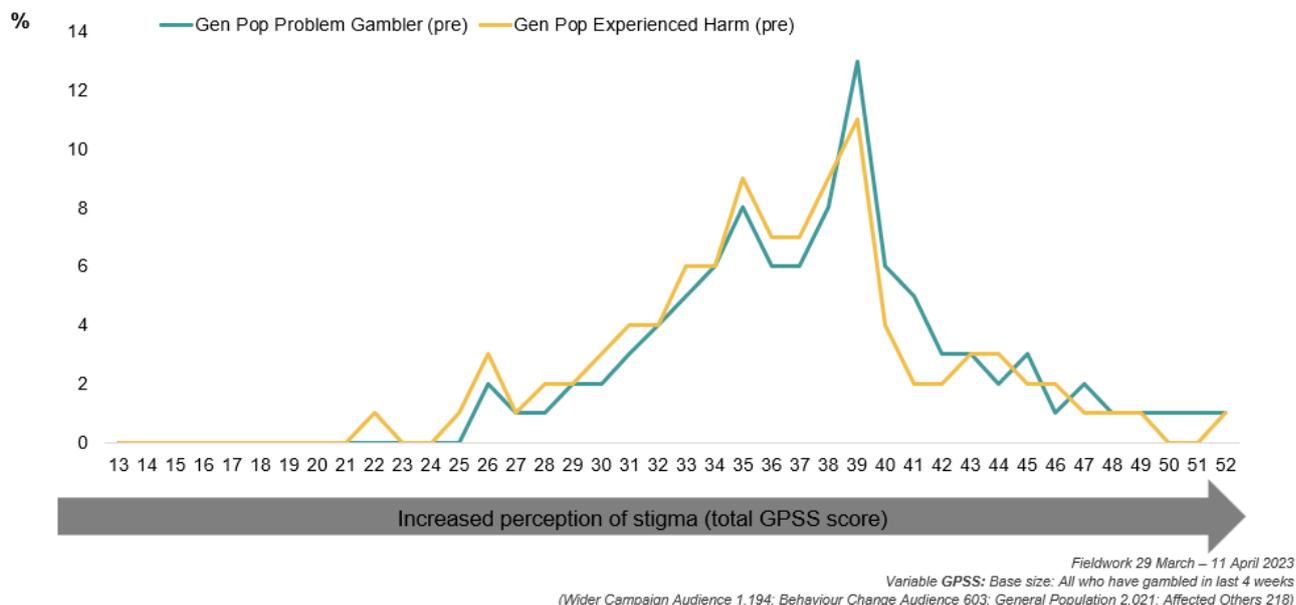
The image below shows the distribution of the GPSS total scores in wave 1. The reductive language typically resulted in higher scores of agreement with the statements (that is, more stigmatising perceptions) than the non-reductive language; however, the pattern of distribution remained similar.

²¹ All participants were shown the following text at the start of the survey to ensure there was shared understanding of what is meant by the term ‘gambling harm’ and similar variations (e.g., “experiencing harms from gambling”): In this survey we will use the term ‘gambling harm’, by which we mean any negative consequences that gambling might lead to in terms of finances, health, quality of life, or wider social issues; both for the people who gamble and the people and communities around them’.

²² The words we choose to describe gambling harms is important. For those who experience harms from gambling, perceived and experienced stigma is a barrier to self-identifying and a barrier to seeking help and support. One of the ways we can act to reduce stigma is through respectful and non-judgmental language. Further guidance and context is provided here in GambleAware’s language guide (2023):

<https://www.begambleaware.org/sites/default/files/2023-04/Stigma%20Language%20guide.pdf>

GPSS: compare individual scores for two different text edits



Source: Ipsos/GambleAware 2023

2.5 Methodology for investigation

The table below provides an overview of the key methods used for investigation. A more detailed description of the full method is available in section 3.

Table 2.2: Investigation aims and approaches

Aim	Approach
To explore the impacts of varying the wording of items in the GPSS using reductive vs non-reductive language.	T-tests were calculated to examine the difference between the two wording version of the questionnaire.
To understand whether all items on the GESS and GPSS contribute equally to the overall measurement of experience and perceived stigma.	Factor analysis was used to better understand the relationship between each item on the scale and the overall scale scores. For GPSS, factor analysis was carried out across the Ostracism and Contempt subscales to investigate whether each item contributed equally to the total score for each subscale.
To identify whether stigma could be classified based on total scores for each scale and investigate the optimum number of classes.	Latent Class Analysis was used to group together total scores into suitable classes. The method explored applying multiple different classes (groups) within the data and compared which were most 'stable'. Discriminant function analysis was then used to assess how accurately each of the items could predict the final classifications.
To consider development of a short-form for each scale.	For the GESS, a short form was developed based on the average mean scores of all items in the scale. Each short-form scale was tested for sensitivity, specificity and balanced accuracy. Different numbers of items were tested (for GESS, 3, 4 and 5-item scales) to find the optimum length. No statistical basis was established for

	<p>developing a short form for the GPSS. However, for the GPSS Ostracism and Contempt subscales, the same method as for the GESS was followed.</p>
<p>To better understand the relationship between GESS and GPSS scores and key demographics.</p>	<p>Regression analysis was carried out on GESS and the two GPSS subscales (Ostracism and Contempt) to reveal the association between demographic characteristics and scale scores.²³ Next, for each scale, chi squared automatic interaction detection (CHAID) was used to segment participants according to the likelihood of having high scores, revealing the demographic characteristics most likely to drive high scores. The results of this is presented in a decision tree.</p>

²³ Note, regression analysis does not explain causality but shows relationships between variables, taking into account any relationship one characteristic may share with another.

3 Key findings

3.1 GESS Key findings

3.1.1 GESS measures a single underlying dimension of experienced stigma, and thus the 'total' score is an appropriate tool for tracking and evaluation.

All items (statements) in the full form GESS contribute similarly to the overall score of stigma; a participant who scores high on one measure tends to score high on most others and a participant scoring low on one question tends to score low on most others. Therefore, a lower score can be relied upon to show a lower measure of experienced stigma regardless of which combination of scores resulted in the total score. This means the total score of the 13 elements in the GESS questionnaire is a meaningful measure of experienced stigma.

3.1.2 There are four natural groupings of experience of stigma as measured by GESS

Statistical analysis shows that allocating total scores into classes is a valid exercise. Having tested applying two, three, four, five or six classes to the data, the optimal number of classes is four, namely: "Low", "Moderate", "High" and "Very high".

Further analysis, using the means and standard deviations of the total score within each class, identified the most suitable cut-off points between each class. Following on from the group showing the least stigma, and by chance, these coincided with increments of 10.

The four classes are distinguished by their differences on the total score produced by summing across the 13 GESS items. The four classes clearly show an increase in the average total GESS score running respectively from GESS averages of 14.6, 25.7, 35.5 and 45.6. The table below shows:

- The average score from the training data in each class.
- The upper and lower cut-off points per class.
- The proportion of the population in each class.²⁴

The robustness of these four groupings was further demonstrated through accurately predicting the class to which each score belonged. These groupings accurately predicted 97% of cases into their correct classes. The model most accurately predicted the low group and least accurately predicted the moderate group.²⁵

²⁴ The reproduced latent classes in the validation subsample show remarkably similar average scores, deviating from their training dataset counterparts by 0.5 or less. Additionally, the pattern of scores for each individual GESS item shows a corresponding pattern of increase across the latent class groups across both training and validation subsamples, with no item standing out as markedly different from any other. Consequently, the LCA quite clearly shows that four classes distinguish between ascending levels of the total score and hold for each item contributing to the total score.

²⁵ 100% of cases that belonged in the low (score of 13-19) group were correctly predicted. 92% of cases that belonged to the moderate (score of 20-29) group were correctly predicted. 7% were incorrectly predicted as belonging to the low group. 96% of cases that belonged to the high (score of 30-39) group were correctly predicted. 5% were incorrectly predicted as belonging to the moderate group. 94% of the very high (score of 40+) group were correctly predicted. 6% were incorrectly predicted as belonging to the high group.

Table 3.1: GESS classification and corresponding scores

	Average score per class	Suggested score boundaries per class	% of population per class
Class 1 Low	14.6	13-19	46.7%
Class 2 Moderate	25.7	20-29	19.9%
Class 3 High	35.5	30-39	21.0%
Class 4 Very high	45.6	40+	12.4%

3.1.3 There may be merit in separately monitoring GESS score of “13”

As a large proportion of participants in the GESS data scored 13 or 14 (that is, “strongly disagree” to all or nearly all statements), the analysis considered whether a five-band classification (splitting low in to two classes) was more appropriate. However, the four-class solution was shown to be more robust.

Unlike the Problem Gambling Severity Index (PGSI), the GESS does not explicitly measure frequency of experience; participants are asked to agree-disagree, rather than never-always. As such, we would argue that a GESS score of 13 should not assume an individual has no experience of stigma.

A five-class model where the “low experience of stigma” class is subdivided (to 13-14 and 15-19) may be desirable and could be explored further; however, this would be a pragmatic rather than statistical solution. Furthermore, a partition to separately monitor a score of 13 or 14 would also make it harder to develop and monitor a short form GESS scale.

3.1.4 A short form of GESS is viable; 3-, 4- and 5-item short forms of GESS are all appropriate tools for campaign evaluation, but trade off accuracy and respondent burden

Analysis shows the short form could comprise as few as three questions, although five questions is suggested as optimal. A 4-item short form balances a manageable number of questions for practical purposes, and has improved accuracy in performance over the 3-item form.

Informed by a Rasch model to help discriminate between items, analysis identified the statements for a 3 and 4-item short form for GESS. The optimal items to include in the 4-item short form are items 3, 4, 6 and 13. The suggested optimal items to include in the 3-item short form are 1, 6 and 13. The 5-item form would include both sets of items: 1, 3, 4, 6 and 13.

Table 3.2: Options for GESS short form scales

Item	3- item scale	4- item scale	5-item scale
1) I feel the need to hide my gambling from my friends	✓		✓
3) Most people would always suspect that I had returned to gambling, even if I did not gamble anymore		✓	✓

4) People have insulted me because of my gambling		✓	✓
6) People can tell that I am a gambler by the way I look	✓	✓	✓
13) Once they know I am a gambler, most people will take my opinion less seriously	✓	✓	✓

There are small but notable differences between the 3-, 4- and 5-item short forms in accurately predicting the correct class for each case. The 3-item short form more accurately predicts the 16-20 (low) class while the four-item short form more accurately predicts the 40+ (very high) class. For the 5-item scale, sensitivity improves for all groups except the “low experience of stigma” group which has 93% accuracy for the 5-item scale and 94% accuracy for the 4-item scale which is a minor difference.

Consideration of a 3-item form

A 3-item form could be considered, to further reduce the burden on the respondent. However, further analysis would be needed to find the optimum combination of the three items. There are 455 possible combinations of three items from the full list of 13 and each of these 455 would need to be tested. Even then, it should be noted that there is no guarantee that any of these would provide a solution that is more accurate in predicting the correct class in each case than the three suggested in this report.

In summary, the 5-item scale is preferred because it performs sufficiently well in classifying across all the groups. It also provides a reasonable balance between accuracy and respondent burden, and includes items that are of interest to wider stigma programme objectives (namely “feeling the need to hide gambling”).

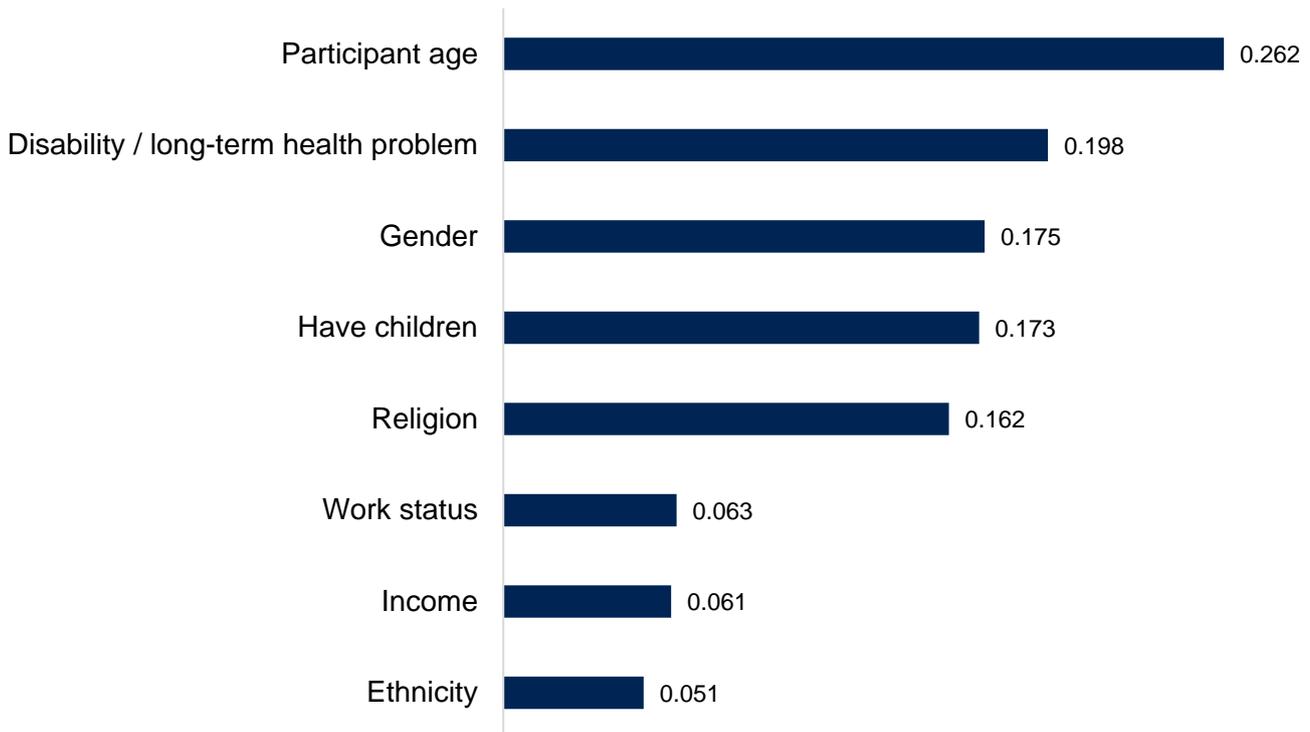
Table 3.3: Predictive accuracy of the short form scales

Scale	Class	Sensitivity	Specificity	Balanced accuracy
5 items	Low (5-7)	0.93	0.97	0.95
	Moderate (8-11)	0.85	0.92	0.88
	High (12-15)	0.8	0.96	0.88
	Very high (16-20)	0.84	0.99	0.91
4 items	Low (4-6)	0.94	0.94	0.94
	Moderate (7-8)	0.77	0.91	0.84
	High (9-12)	0.72	0.96	0.84
	Very high (13-16)	0.8	0.99	0.89
3 items	Low (3-4)	0.9	0.97	0.93
	Moderate (5-6)	0.72	0.91	0.82
	High (7-9)	0.83	0.91	0.87
	Very high (10-12)	0.74	0.99	0.86

3.1.5 Experience of stigma is greatest among young religious males

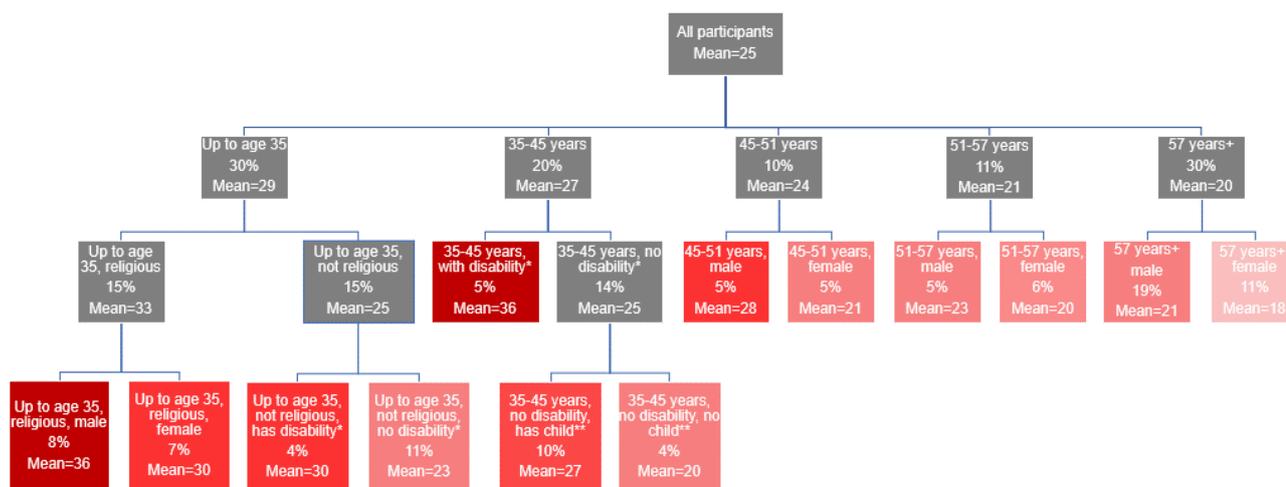
Analysis of participants by social-economic characteristics showed some groups are more likely to have a high to very high experience of stigma. Regression analysis shows that 28% of the variation in GESS scores can be explained by differences in demographics. The strongest predictors of GESS scores are age, followed by disability and gender, whilst religion and presence of children have a similar effect on GESS scores (adjusted for other variables in the model).

Figure 3.1: Key Drivers for GESS



CHAID analysis was conducted to show how demographics interact and potentially compound the risk of experiencing stigma, identifying which segments of the population have the highest GESS scores. As with regression results, the most discriminating demographic variable is age, followed by disability, gender and religion. The segment group with the highest GESS score is comprised of under-35 religious males (8.1% of the sample), followed by middle aged respondents with a disability. Conversely, the segment groups least likely to have high GESS scores are elderly females, or middle-aged people who are not disabled and don't have children.

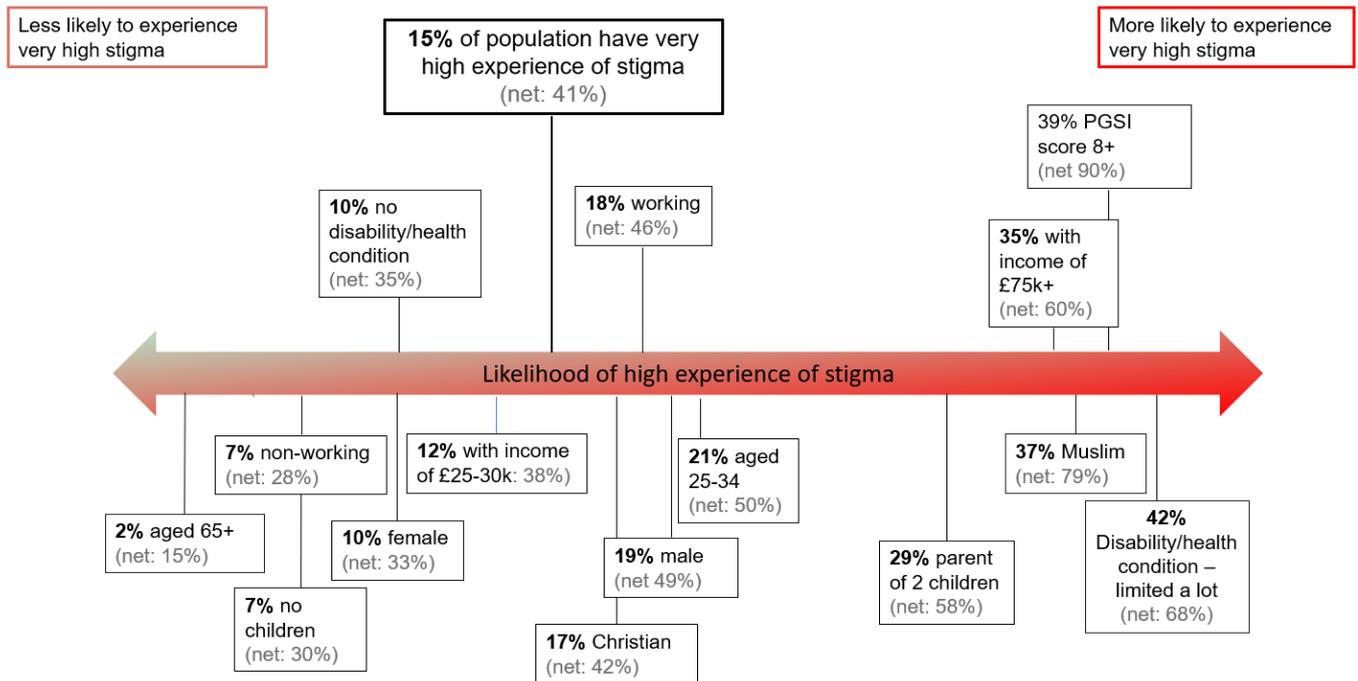
Figure 3.2: GESS CHAID analysis



*Has disability or long-term health condition.
 **At least one child in household
 Mean=Mean GESS score

The image below provides a visual map of key differences in high experience of stigma (i.e. a score of 30+ using the GESS). Very high experience of stigma is particularly marked for some specific groups: those with a severely limiting disability or health condition, those of Muslim faith, and individuals with an income of £75,000+. This is unsurprising to some extent due to intersectional stigma among marginalised communities. However, the researchers were surprised about level of income given that people with lower socio-economic status are more likely to experience gambling harms. However, regression analysis showed that although there is a correlation between high income and high experience of stigma, this correlation can be explained by having children in the household as those with children are more likely to be higher earners.

Figure 3.3: Demographic characteristics and high experience of stigma



% in bold: % with very high experience of stigma
 % in grey: net % with very high/high experience of stigma

Further analysis using regression modelling shows the key demographic drivers for GESS, with age being the biggest driver, followed by disability, gender and the presence of children in the household. These are discussed further below.

In summary;

- **Age** is the biggest discriminant of GESS. Younger age groups (18 to 44 years) are more likely to have a very high experience of stigma than older age groups (45+).
- Having a **disability** is the second strongest predictor of GESS score. People who have a disability that limits their day-to-day life a lot are more likely (42%) to have a very high experience of stigma than those without a disability at all.
- **Gender** is the third strongest predictor of GESS score. Men (19%) are more likely than women (10%) to experience very high stigma.
- **High income** participants are more likely to have a very high experience of stigma than moderate- and low-income participants. However, this is likely influenced by the presence of children in the household; high income itself is not a strong driver of GESS score once the model has controlled for other variables.
- **Work status** is not a strong predictor of GESS score. Participants who said they were working appear more likely to have a very high experience of stigma than those who said they were not working; however, this variation is explained by age (older age groups are less likely to be in work).
- **Religion** is a notable but less strong predictor of GESS score. People who practice a religion (particularly Muslim or Jewish) are more likely than those who do not to experience very high stigma.

- Having **one or more child** is a similarly strong predictor of GESS score. People with children in the household are more likely than those without to have very high experience of stigma.
- **Geographical region** has a negligible effect on GESS.

3.1.6 There is a close relationship between socio-demographic predictors of GESS and socio-demographic predictors of PGSI.

The higher a participant's **PGSI score**, the more likely they are to have a very high experience of stigma. The difference is most notable between those with a PGSI score of 3-7 (9% have very high experience of stigma) and a score of 8+ (39% have very high experience). Of those with a PGSI score of 0, 1% experience very high stigma; 4% of those with a PGSI score of 1-2 experience very high stigma. Moreover, when added to a regression analysis, PGSI predicts 60% of the variation in GESS scores.

When controlling for PGSI, household income, ethnicity, disability and working status no longer account for any further variation in GESS. Though age, gender, religion and children are still statistically significant, they only account for a further 1.3% of variation in GESS once PGSI has been controlled for. This suggests that the same set of socio-demographics predict both GESS and PGSI.

It should be noted that the relationship between PGSI and socio-demographics is complex, and will not be universal across all variables. For example, whilst PGSI will not have an impact on age, ethnicity and gender, it may influence other more fluid demographics such as income, marital status, working status. Further multi-variate regression may help better understand the relationship between PGSI, demographics and GESS, where PGSI and GESS are considered as two dependent variables.

3.2 GPSS Key findings

3.2.1 Replacing reductive with non-reductive language in the question wording of the 13 items has minimal impact on measuring perceived stigma.

Waves 1 and 2 of the GambleAware Stigma Campaign Evaluation used a split sample to test two alternative wordings for the GPSS. One used the original text, referring to “problem gamblers”, the second version used a revised non-reductive text referring to “people experiencing harms from gambling”.

Both versions of question wording were found to have similar structural integrity: i.e., respondents gave similar answers, and the individual items combined together in a similar way to create an overall assessment of perceived stigma. The use of reductive language tended to receive, on average, slightly higher scores than non-reductive language. These changes per item were significant but small and, for some items, there was no difference. Despite the statistical significance of most of the wording differences, our expectation is that these generally are too small to have a major impact on further

analysis. There is evidence of an accumulative effect of the differences for each item on the total score but this is also very small.²⁶

Table 3.4: Comparison of summary scores for original (reductive) and revised (non-reductive) text

	Original text: “Problem gambler”	Revised text: “People experiencing harms from gambling”
Summary Score: Contempt	16.76	16.13
Summary Score: Ostracism	21.09	20.61

We therefore recommend that future use of the GPSS scale uses non-reductive language. This will ensure that the monitoring of GPSS does not undermine GambleAware’s wider commitment to use appropriate language that helps reduce stigma.

3.2.2 The GPSS provides greater value when Ostracism and Contempt are analysed independently as two subscales.

Factor analysis confirmed previous research that GPSS measures two dimensions of perceived stigma: Ostracism and Contempt. Analysis proved a high degree of consistency between the items.²⁷

Whilst the Ostracism and Contempt subscales largely tend to reproduce the total GPSS score distribution, and there is a strong correlation between each subscale and the total GPSS score, Ostracism and Contempt correlate less highly with each other.²⁸ There is not a clear correlation between scores on items from one scale with scores on items from the other scale; not all items cleanly measured either Ostracism or Contempt, some items measured aspects of both GPSS subscales.

This was confirmed in subsequent analysis which also compared natural groupings (“classes”) within each subscale. This demonstrated that there is not a strong correlation between each class on the Contempt score and the equivalent class on the Ostracism score; while people who score high in one subscale are likely to score high in the other subscale, a significant minority of people may score high on one scale but low on the other.

As shown in the table below, a substantial minority of people have slightly mixed perceptions towards Ostracism and Contempt. Consequently, perceptions of these two aspects of stigma do not necessarily move in the same direction and should be considered separately. We therefore recommend that moving forwards, GambleAware should continue to monitor trends in Ostracism and Contempt independently of each other, rather than monitoring change in the GPSS score overall.

²⁶ Given the differences in impact of the variation in wording, the combined data from both samples was used for the remainder of the analysis.

²⁷ Cronbach’s alpha measures for the internal consistency of responses were 0.82 for Contempt and 0.86 for Ostracism. These are both highly respectable scores indicating a high degree of consistency between the items.

²⁸ The correlation between the three scales shows that Ostracism correlates $r = 0.92$ with the GPSS total, and the corresponding $r = 0.94$ for Contempt. In other words, there is a close correspondence between a person’s level of perceived Ostracism and a person’s level of perceived Contempt and the total GPSS score. However, Ostracism and Contempt correlate less highly with each other, $r = 0.71$

Table 3.5: Distribution of Ostracism scores by Contempt groupings

Contempt	Ostracism			
	6-15	16-20	21-24	N
7-16	59%	38%	3%	717 (row percentage)
17-21	15%	76%	9%	2330
22-28	3%	44%	53%	958

3.2.3 A three-class model for each subscale within GPSS is optimal.

Latent Class Analysis showed that classifying total scores into smaller groups is a valid exercise, and three and four-class models were tested for each subscale. The three-class model performed better, with the four-class model being unstable.

A key finding is the tendency for few people to score “2” or under (“slightly disagree”) across all individual items. In general, many people seem to at least strongly agree to many of the items, resulting in a large number of people falling into the intermediate class. As such, there is little scope for further splitting the “moderate” group.

Further analysis identified the cut-off points between each class. Next, the average scores for each subgroup shows an increase of 4 points for the Ostracism subscale and 5 points for the Contempt subscale.

Table 3.6: Ostracism subscale classification

	Average score per class	Suggested score boundaries per class	% of population per class
Class 1 Low	13.3	6-15	20%
Class 2 Moderate	17.8	16-20	62%
Class 3 High	21.9	21-24	18%

Table 3.7: Contempt subscale classification

	Average score per class	Suggested score boundaries per class	% of population per class
Class 1 Low	14.6	7-16	18%
Class 2 Moderate	19.4	17-21	58%
Class 3 High	24.3	22-28	24%

3.2.4 Results of the analysis do not support the proposal of a short-form for the GPSS

The creation of a short-form GPSS trades off respondent's burden with quality and accuracy. Short-form versions of the Contempt and Ostracism scales do provide some reasonable accuracy; however, they are not advised due to a significant level of misclassifications.

To provide a short form solution that offers significant advantage over the long forms, short forms comprising three questions each were tested for each subscale, which would result in a total GPSS short form of 6 items instead of 13.

To test a short form for each subscale, the same process was followed. The three items most likely to predict the total score were selected. Analysis of the total score for those items suggested suitable ranges for each class: low, moderate, high. The items in the shortform were tested to see how well they performed in accurately allocating the total score into the correct class (in comparison to the accuracy of the longform). Accuracy was measured by reviewing sensitivity and specificity.²⁹ The average of these provides an assessment of balanced accuracy for how well the form has allocated total scores to each of the three classes.

Ostracism short form

For Ostracism a three-item shortform results in a substantial number of misclassifications, most frequently for the low and moderate classes.³⁰ Although different cut-off points were tested, there were a substantial number of misclassifications. Overall, this short form grouping reproduced 86% of the long form classification.

²⁹ Sensitivity looks at where cases have been allocated to a specific class and checks whether that allocation was correct (that is, how many false positives there were). Specificity looks at where cases have not been allocated to a specific class and checks whether that was correct (that is, how many false negatives there were.)

³⁰ Shortform items chosen were:

'Many people would be uncomfortable communicating with a problem gambler/person experiencing harms from gambling'

'Most people would not hire a problem gambler/person experiencing harms from gambling'

'Most people would not want to enter into a committed relationship with someone they knew problem gambler/person experiencing harms from gambling'

Table 3.8: Accuracy of Ostracism short form by class

Class	Sensitivity	Specificity	Balanced accuracy
Low (3-7)	0.73	0.98	0.85
Moderate (8-10)	0.91	0.77	0.84
High (11-12)	0.82	0.95	0.89

Contempt short form

The Contempt short form scale performs reasonably well in terms of sensitivity for the moderate group, with specificity high for low and high groups; however, again there are substantial misclassifications. Overall, this short form grouping correctly reproduced 82% of the long form classification.³¹

Table 3.9: Accuracy of Contempt short form by class

Class	Sensitivity	Specificity	Balanced accuracy
Low (3-7)	0.5	0.98	0.77
Moderate (8-9)	0.93	0.65	0.79
High (10-12)	0.72	0.97	0.84

3.2.5 Socio-demographics are not a strong predictor of perceived stigma.

Although some socio-economic groups are more likely to show high perceived gambling stigma than others, regression analysis shows that socio-demographic characteristics were, at best, only weakly related to Contempt and Ostracism scores – demographics included within the analysis explained only 4% of the variance within Contempt and 1% of the variance for Ostracism. It is likely that differences between demographics subgroups can be explained by other factors, such as personal attitudes, values and experience.

Differences seen in both Ostracism and Contempt subscale scores:

³¹ Short-form items chosen:

Most people think problem gamblers/people experiencing harms from gambling are lazy.

Once they know a person is a problem gambler/person experiencing harms from gambling, most people will take his or her opinion less seriously.

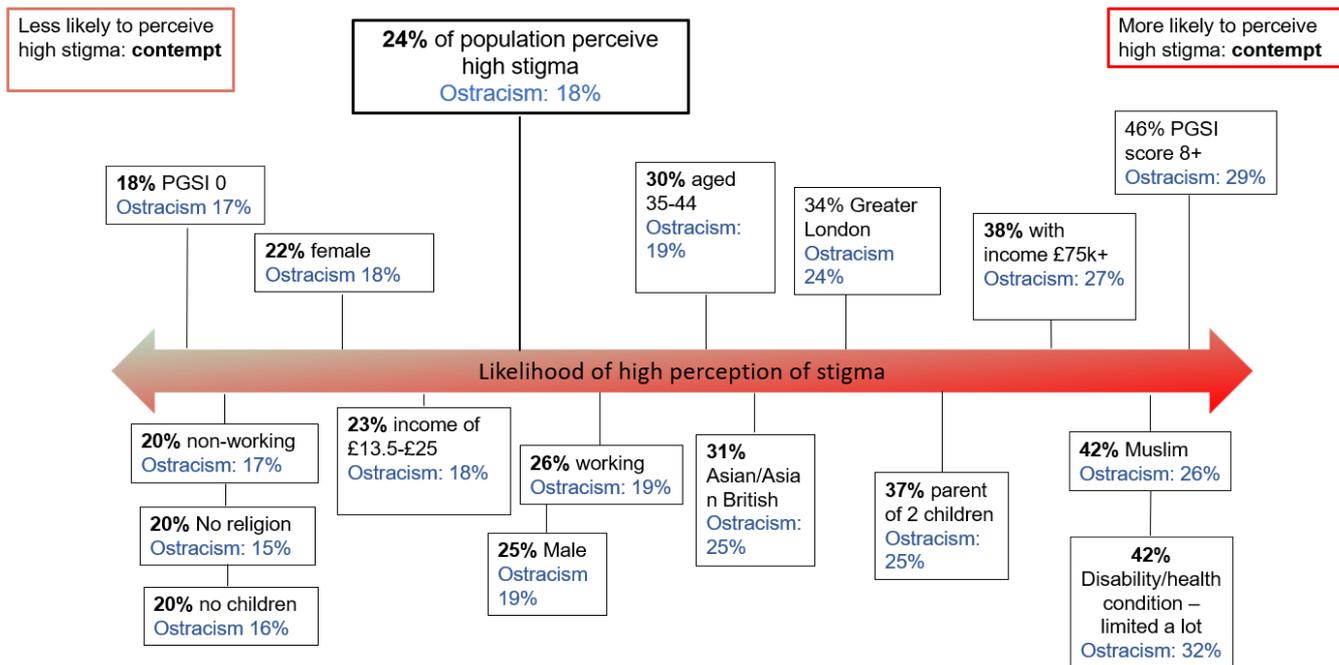
Most people believe that problem gamblers/people experiencing harms from gambling have no self-control.

- Those with a high household **income** (£75k+ per year) are more likely than those in lower incomes groups to show high perceived gambling stigma.
- Those with **children** (particularly two children) are more likely than those without children to show perceived gambling stigma.
- People whose day-to-day lives are limited a lot by **disability** or long-term health conditions are more likely to show high perceived gambling stigma.
- People holding a **religion** (particular Muslims, Jewish people and, to a lesser extent, Hindus) are more likely to show high perceived gambling stigma than those without a religion.
- Those with a **PGSI** score of 8+ are more likely to show high perceived gambling stigma than those with a lower PGSI score.

Differences seen most notably in Contempt subscale scores:

- **Males** are more likely to show high perceived gambling stigma than females.
- Younger **age** groups (under 45) are more likely to show high perceived gambling stigma than older age groups (there is little variation by age in Contempt scores but those aged 35-44 are more likely to show higher perceived gambling stigma).
- Non-White **ethnic groups** are more likely to show high perceived gambling stigma than those with white ethnic background (in Contempt scores, Black and Asian groups are more likely to score higher than other ethnic groups).
- People who are in **work** are more likely to show high perceived gambling stigma than those who are not working.

Demographic characteristics and high perception of stigma



% in bold: % with high perception of stigma - contempt
% in blue: % with high perception of stigma - ostracism

4 GESS Technical Annex

4.1 Assessing structural integrity

There is strong evidence to support the existence of a single underlying dimension of experienced stigma, arising from the factor analysis and internal consistency analysis run on the data. The first stage of the analysis provided evidence that the 13 GESS items comprised a single scale using factor analysis, which is a statistical technique that reveals the number of dimensions underlying responses to the 13 items. The results for both the training and validation sample both resulted in a single factor extraction, confirming that the items represented a single underlying dimension. Table 3.1 also shows the factor loadings for each item on the scale separately for the training and validation subsamples. In both cases, all loadings are high and show little variation between them, indicating that each item has an approximately equivalent relationship to the underlying dimension of experienced stigma.

Having confirmed that a single underlying dimension was being measured by the GESS items, a Cronbach's alpha measure of the "internal consistency" of responses to the items showed a high degree of internal consistency, showing that people treated the items as belonging to a single underlying dimension of experienced stigma. Cronbach's alpha varies between 0 and 1, with one being the highest measure of consistency, i.e., people scoring high on one item will also tend to score high on all other items. The evidence shows very high levels of internal consistency for both the training subsample, alpha = 0.972 and for the validation dataset alpha = 0.973. Removal of any item would not improve the level of alpha, as indicated by the columns headed 'Alpha if item deleted' in Table 3.1, confirming the evidence from approximately factor loadings. Additionally, the column head R² shows how well each item is predicted from the other items in the scale (R² has a maximum value of one, indicating perfect prediction), again showing a high degree of consistency between the items.

Table 4.1: Results of factor analysis and reliability tests for GESS

<i>Question item</i>	Factor loadings		R²		Alpha if item deleted	
	<i>Train</i>	<i>Valid</i>	<i>Train</i>	<i>Valid</i>	<i>Train</i>	<i>Valid</i>
I feel the need to hide my gambling from my friends	0.883	0.885	0.751	0.757	0.969	0.970
I sometimes have the thought that I have screwed up my life by gambling	0.894	0.904	0.774	0.790	0.969	0.970
Most people would always suspect that I had returned to gambling, even if I did not gamble anymore	0.831	0.837	0.656	0.665	0.971	0.971

People have insulted me because of my gambling	0.884	0.878	0.757	0.744	0.969	0.970
I have the thought that I should be ashamed of myself for my gambling	0.879	0.893	0.742	0.772	0.969	0.970
People can tell that I am a gambler by the way I look	0.854	0.822	0.714	0.643	0.970	0.972
Others think I am not worth the investment of time and resources because I am a gambler	0.881	0.890	0.744	0.766	0.969	0.970
I sometimes have the thought that I deserve the bad things that have happened to me in life because I gamble	0.887	0.890	0.753	0.762	0.969	0.970
I feel the stress in my life is what causes me to gamble	0.848	0.858	0.686	0.708	0.970	0.971
Others view me as morally weak because I am a gambler	0.870	0.898	0.724	0.775	0.970	0.970
I avoid situations where another person might have to depend on me	0.840	0.810	0.664	0.619	0.970	0.972
I don't think I can be trusted because I gamble	0.878	0.881	0.742	0.754	0.969	0.970
Once they know I am a gambler, most people will take my opinion less seriously	0.823	0.834	0.640	0.665	0.971	0.971

Note: training n = 1,287; validation n = 1,272.

4.2 Grouping the summary score

Latent Class Analysis (LCA) is a statistical technique that groups together cases with similar patterns of responses across a set of question items. There is no guarantee that LCA will produce a pattern of results that is based on grouping cases according to the value of the summary score. However, given the findings of the reliability and factor analysis, which suggest a very high degree of consistency between the items, a grouping based on the summary score is anticipated rather than one which produces commonalities between different subsets of items.

We explored LCA solutions using between two and six latent classes on the training dataset and found that a four-class solution provided the most viable model. This model showed a stable structure when reproduced on the validation subsample and was chosen as the preferred solution. The four classes are

distinguished by their differences on the total score produced by summing across the 13 GESS items. The four classes clearly show an increase in the average total GESS score running respectively from GESS averages of 14.6, 25.7, 35.5 and 45.6 in the training subsample. The reproduced latent classes in the validation subsample show remarkably similar average scores, deviating from their training dataset counterparts by 0.5 or less. Additionally, the pattern of scores for each individual GESS item shows a corresponding pattern of increase across the latent class groups across both training and validation subsamples, with no item standing out as markedly different from any other. Consequently, the LCA quite clearly shows that four classes distinguish between ascending levels of the total score and hold for each item contributing to the total score.

Table 4.2: GESS item and total scores for latent class groups

Question Item	Training				Validation			
	Class 1 <i>("Low experience of stigma")</i>	Class 2 <i>("Moderate experience of stigma")</i>	Class 3 <i>("High experience of stigma")</i>	Class 4 <i>("Very high experience of stigma")</i>	Class 1 <i>("Low experience of stigma")</i>	Class 2 <i>("Moderate experience of stigma")</i>	Class 3 <i>("High experience of stigma")</i>	Class 4 <i>("Very high experience of stigma")</i>
<i>GESS Total Score Average</i>	14.6	25.7	35.5	45.6	14.4	25.3	35.0	45.1
I feel the need to hide my gambling from my friends	1.1	2.0	2.8	3.6	1.1	2.0	2.8	3.5
I sometimes have the thought that I have screwed up my life by gambling	1.0	1.9	2.8	3.5	1.0	1.8	2.8	3.6
Most people would always suspect that I had returned to gambling, even if I did not gamble anymore	1.3	2.1	2.9	3.6	1.2	2.0	2.8	3.5
People have insulted me because of my gambling	1.0	1.8	2.7	3.4	1.0	1.7	2.6	3.5
I have the thought that I should be ashamed of myself for my gambling	1.1	2.1	2.8	3.6	1.1	2.0	2.8	3.5

People can tell that I am a gambler by the way I look	1.0	1.8	2.5	3.4	1.0	1.7	2.4	3.3
Others think I am not worth the investment of time and resources because I am a gambler	1.1	2.0	2.7	3.5	1.1	2.0	2.7	3.5
I sometimes have the thought that I deserve the bad things that have happened to me in life because I gamble	1.1	1.9	2.7	3.5	1.1	2.0	2.7	3.4
I feel the stress in my life is what causes me to gamble	1.1	2.1	2.7	3.5	1.1	2.1	2.7	3.5
Others view me as morally weak because I am a gambler	1.1	2.0	2.7	3.5	1.1	2.0	2.7	3.5
I avoid situations where another person might have to depend on me	1.2	2.0	2.8	3.6	1.2	2.0	2.7	3.4
I don't think I can be trusted because I gamble	1.1	1.8	2.7	3.3	1.0	1.8	2.6	3.4
Once they know I am a gambler, most people will take my opinion less seriously	1.3	2.2	2.8	3.6	1.3	2.2	2.7	3.5

Note: Respondents with gambling experience, training n = 1,287; validation n = 1,272.

Having established that the LCA produces a set of groups distinguishing people by their total score across the individual 13 items, a practical approach was sought to create these groupings through assigning cut-off points to the total score. The alternative of applying the model results to the items each time a new survey is carried out was deemed overly cumbersome. Consequently, using the means and standard deviations of the total score within each class, an initial set of cut-off scores were identified and used to create a classification. These groups were then cross-tabulated with the latent class membership to assess the degree of overlap between the “eyeball” approach and the LCA solution. Alternative cut-off points were then made based on the extent of misclassification and this procedure repeated until the best overlap was found. By chance, the most appropriate groupings coincided with increments of 10 from the group showing least stigma, i.e. 13-19, 20-29, 30-39, 40+.

A comparison between the groups created directly using total score cut-off points and the latent classes showed high levels of correspondence on both the training and, more importantly, the validation

subsamples. For the validation sample, the correspondence between each GESS total score group and the corresponding latent class was between 96% and 100%, which is substantially better than the correspondence between the two groupings seen for the training sample.

Table 4.3: GESS total score and latent class group correspondence

	Grouped score	Latent Classes				Total n	%
		Low	Moderate	High	Very high		
Training	<20	98.0%	2.0%	0.0%	0.0%	601	46.7%
	20-29	10.5%	85.9%	3.5%	0.0%	256	19.9%
	30-39	0.0%	12.2%	85.9%	1.9%	270	21.0%
	40+	0.0%	0.0%	11.9%	88.1%	160	12.4%
Validation	<20	99.7%	0.3%	0.0%	0.0%	601	47.2%
	20-29	0.8%	96.7%	2.5%	0.0%	241	18.9%
	30-39	0.0%	4.1%	95.5%	0.3%	290	22.8%
	40+	0.0%	0.0%	4.3%	95.7%	140	11.0%

A discriminant function analysis, using the validation subsample, included the 13 GESS items to predict the group membership based on the total score cut-off points. Overall, the model accurately predicted around 97% of cases into their actual groups. The under 20 group was most accurately predicted at 100% and the 20-29 group least accurately predicted at 92%, with seven per cent misclassified as belonging to the under 20 group. Around 96% of the 30-39 group were accurately predicted with five per cent misclassified as 20-29. For the 40+ group, 94 per cent were accurately predicted with six per cent misclassified as 30-39. 96.7% of original grouped cases correctly classified.

Table 4.4: Predicting the cut-off groups from the GESS items

	Original Group	Predicted Group Membership				Total
		<20	20-29	30-39	40+	
Count	<20	601	0	0	0	601
	20-29	16	221	4	0	241
	30-39	0	13	277	0	290
	40+	0	0	9	131	140
%	<20	100.0	0.0	0.0	0.0	100.0
	20-29	6.6	91.7	1.7	0.0	100.0
	30-39	0.0	4.5	95.5	0.0	100.0
	40+	0.0	0.0	6.4	93.6	100.0

Given the high degree of correspondence between the groups based on the total score cut-off points and the latent class groups, the profile of the cut-off groups, in terms of breakdowns of total and item scores, was very similar to that of the LCA groups (Table 3.5).

Table 4.5: GESS item and total scores by net groups

Question Item	Training				Validation			
	<20	20-29	30-39	40+	<20	20-29	30-39	40+
Total GESS score average	14.4	25.1	34.8	45.3	14.4	25.2	34.9	45.0
I feel the need to hide my gambling from my friends	1.1	2.0	2.7	3.5	1.1	2.0	2.8	3.5
I sometimes have the thought that I have screwed up my life by gambling	1.0	1.9	2.7	3.5	1.0	1.8	2.7	3.6
Most people would always suspect that I had returned to gambling, even if I did not gamble anymore	1.2	2.1	2.8	3.6	1.2	2.0	2.8	3.5
People have insulted me because of my gambling	1.0	1.7	2.6	3.4	1.0	1.7	2.6	3.4
I have the thought that I should be ashamed of myself for my gambling	1.1	2.0	2.7	3.6	1.1	2.0	2.8	3.5
People can tell that I am a gambler by the way I look	1.0	1.7	2.5	3.4	1.0	1.7	2.4	3.3
Others think I am not worth the investment of time and resources because I am a gambler	1.1	1.9	2.7	3.5	1.1	2.0	2.7	3.4
I sometimes have the thought that I deserve the bad things that have happened to me in life because I gamble	1.0	1.9	2.7	3.4	1.0	2.0	2.7	3.5
I feel the stress in my life is what causes me to gamble	1.1	2.0	2.7	3.4	1.1	2.0	2.7	3.4
Others view me as morally weak because I am a gambler	1.1	2.0	2.6	3.5	1.1	2.0	2.7	3.5
I avoid situations where another person might have to depend on me	1.2	2.0	2.8	3.5	1.2	2.0	2.6	3.4

I don't think I can be trusted because I gamble	1.0	1.8	2.6	3.4	1.0	1.8	2.6	3.4
Once they know I am a gambler, most people will take my opinion less seriously	1.2	2.2	2.7	3.5	1.3	2.1	2.8	3.5

4.3 Short-form development

4.3.1 Overview

The use of 13 items to create a scale may be unnecessarily burdensome on respondents if a smaller number of items can serve the same purpose as the full set of 13 items. At one extreme, only one item would be required because everybody gave exactly the same response to that item, i.e., all items are perfectly correlated. However, we know that is not the case with the GESS data. Yet, from Table 2.1, the results from the factor and reliability analyses show that the items are all strongly related to each other. Consequently, it should be possible to achieve a similar spread of people across the experienced stigma dimension using a relatively small number of items, i.e., to create a short-form version of the complete (long-form) version of the scale.

When considering a short-form scale there are both practical and technical considerations to address. We know, in advance, that the predictive capability of short-form is unlikely to be as high as the long-form, but the location and size of the prediction errors may be more or less important depending upon the uses intended for the scale. Our understanding is that the primary aim of the scale is to monitor change over time rather than for use as a diagnostic tool. Consequently, it is probably appropriate to aim for a solution that balances prediction errors across the range of the long form rather than trying to maximise accuracy for any one endpoint of the scale, as may be the case where, for example, the intention is to maximise the accuracy of detecting people experiencing extreme forms of stigma.

A key decision is the number of items to select and then how to assign scores on the restricted range of the short-form (e.g. with 3-items, the range is between 3 and 12) to the range of the long-form (a range of 13 to 52). Working with the grouped scores facilitates this process because we can use the four groups from the long-form as the “gold standard” and assess how best to combine the scores on the short-form to correspond to the long-form groups.

In general, the greater the number of items on the short-form, the easier it becomes to mirror the long-form distribution from the short-form. However, our preference is for a more parsimonious number of items for the short-form. We consider short-forms comprising 3-, 4- and 5-items.

4.3.2 Measuring accuracy of the short form scales

For the purpose of creating a short-form scale, “accuracy” refers to making a correct prediction of the long-form group from the corresponding short-form group and correctly predicting which respondents do not belong in the group. Both predictions can be accurate (true) and inaccurate (false). If we label the in-group prediction as positive and the out-of-group prediction as negative, we can distinguish four states, as described in Table 4.6 below. We note that each of the four groups of the GESS grouped scores has its own set of prediction statistics for accuracy.

Table 4.6: Classifying accurate and inaccurate predictions of the GESS long form from the short form

	True Score	
Predicted Score	True	False
True	True positive (TP): Correctly predicted in group.	False positive (FP): Incorrectly predicted in-group when it is actually out of group.
False	False negative (FN): Incorrectly predicted out-of-group when is actually in-group.	True negative (TN): Correctly predicted out of group.

From these combinations of true and false positives and negatives, we discuss the following measures of accuracy:

- **Sensitivity:** what proportion of those in the long form group are correctly predicted by the short form (aka the true positive rate, TPR)?
- **Specificity:** what proportion of those not in the group on the long form are predicted not to be in the group by the short form (aka the true negative rate, TNR)?
- **Balanced Accuracy:** the average of sensitivity and specificity.

4.3.3 Number of short-form items

We considered different possible short-forms combinations with either 3-, 4- or 5-items. A key aim in selecting the items was to try to ensure a good range across the 13-item summary score distribution. In other words, we wanted to select three items with a low, a high and a medium average score. Item “13” had the highest mean score of 2.03, whereas item 6 had the lowest mean score of 1.73 and item “1” was around the middle with an average score of 1.91 (Table 4.7). For the 4-item scale we chose two questions on the extremes and two either side of the mid-point. For the 5-item scale all five of the questions identified for the 3- and 4-item scales were included.

Table 4.7: Average scores to the 13 GESS questions

Item	Mean
I feel the need to hide my gambling from my friends	1.91
I sometimes have the thought that I’ve screwed up my life by gambling	1.86

Most people would always suspect that I'd returned to gambling, even if I didn't gamble anymore	2.00
People have insulted me because of my gambling	1.79
I have the thought that I should be ashamed of myself for my gambling	1.93
People can tell that I am a gambler by the way I look	1.73
Others think I am not worth the investment of time and resources because I am a gambler	1.89
I sometimes have the thought that I deserve the bad things that have happened to me in life because I gamble	1.85
I feel the stress in my life is what causes me to gamble	1.91
Others view me as morally weak because I am a gambler	1.89
I avoid situations where another person might have to depend on me	1.97
I don't think I can be trusted because I gamble	1.80
Once they know I am a gambler, most people will take my opinion less seriously	2.03

Table 4.8: Questions used in short-form development

Question	3- item scale	4- item scale	5- item scale
1) I feel the need to hide my gambling from my friends	Yes		Yes
3) Most people would always suspect that I had returned to gambling, even if I did not gamble anymore		Yes	Yes
4) People have insulted me because of my gambling		Yes	Yes
6) People can tell that I am a gambler by the way I look	Yes	Yes	Yes
13) Once they know I am a gambler, most people will take my opinion less seriously	Yes	Yes	Yes

4.3.4 Grouping the short-form scores

The challenge in grouping the short-form 3-item scale accurately is immediately apparent when we consider the cut-off point for the lowest score group in the long-form (Table 2.8). Grouping together the first two scores in the short-form ("3" and "4") would miss a substantial 104 respondents who scored '5' in the short-form (along with the 19 scoring "6" and "7"). Conversely, increasing the short-form bottom score group to cover scores 3-5, would result in the incorrect inclusion of 115 respondents who belong to the long-form 20-29 group. There is no obvious solution to this challenge, except possibly to extend the number of items and see if that offers an improvement. Alternatively, we could try all possible combinations of 3 items from the 13, which is 455 3-item scales to explore an improvement. However, there is no guarantee that any of these solutions would provide an appropriate solution.

The information provided in Tables 2.9-2.11 was used to create separate 3-, 4- and 5-item scale threshold scores as cut-off points for the four groups using the scores shown in Table 2.12, which provides accuracy performance statistics for the scales.

Table 4.9: Short-form score on 3-item scale

3-item short form	Long-form group				Total
	<20	20-29	30-39	40+	
3.00	897	8	0	0	905
4.00	182	30	2	0	214
5.00	104	115	9	1	229
6.00	18	245	59	0	322
7.00	1	71	117	3	192
8.00	0	24	173	12	209
9.00	0	4	172	63	239
10.00	0	0	26	73	99
11.00	0	0	2	83	85
12.00	0	0	0	65	65
Total	1202	497	560	300	2559

Table 4.10: Short-form score on 4-item scale

4-item short form	Long-form group				Total
	<20	20-29	30-39	40+	
4.00	813	7	0	0	820
5.00	194	20	1	0	215
6.00	128	59	1	0	188
7.00	51	96	9	0	156
8.00	15	221	36	0	272
9.00	1	66	80	1	148
10.00	0	25	125	6	156
11.00	0	3	147	9	159
12.00	0	0	131	45	176
13.00	0	0	24	69	93
14.00	0	0	4	63	67
15.00	0	0	2	59	61
16.00	0	0	0	48	48
Total	1202	497	560	300	2559

Table 4.11: Short-form score on 5-item scale

	Long-form group	Total
--	-----------------	-------

5 item short-form	<20	20-29	30-39	40+	
5.00	791	3	0	0	794
6.00	178	4	0	0	182
7.00	145	30	0	0	175
8.00	61	63	1	0	125
9.00	25	88	5	0	118
10.00	2	201	14	0	217
11.00	0	69	60	0	129
12.00	0	29	94	1	124
13.00	0	8	116	3	127
14.00	0	2	119	8	129
15.00	0	0	119	37	156
16.00	0	0	25	50	75
17.00	0	0	6	63	69
18.00	0	0	1	54	55
19.00	0	0	0	45	45
20.00	0	0	0	39	39
Total	1202	497	560	300	2559

4.3.5 Predictive accuracy of the subscales

In terms of overall accuracy across all cells of the tables, there was little to choose between the three scales with the 3-item scale scoring 0.83, the 4-item scale scoring 0.84 and the 5-item scale scoring 0.87. As expected, all three versions of the short form scales were better at accurately detecting people who were outside the target group of interest (true negatives), i.e., specificity was generally high and always above 0.9.

More challenging for the short form scales is capturing those people who actually belong to a group, as defined by their long form score. The 3-item scale performs least well in this regard for the next-to-bottom stigma group, capturing 72% of people who should be in the group (sensitivity) and just under three-quarters (74%) of those in the top group. The 4-item scale improves the sensitivity performance of all the groups except the next-to-top group which captures 83% of the target group in the 3-item scale and 72% in the 4-item scale. For the 5-item scale, sensitivity improves for all groups except the bottom group which captures 93% of the target group for the 5-item scale and 94% for the 4-item scale, a minor difference.

In summary, the 5-item scale is preferred because it performs reasonably well in classifying across all the groups. The 4-item scale also provides good classification for the top and bottom groups but performs less well for the two intermediate score groups. The 3-item scale is good at predicting the bottom group and the next-to-top group but less good with the remaining groups. If the interest lies in monitoring change around the top group, e.g., monitoring trends and exploring potential intervention impacts, then the 3-item scale will have more classification issues with moves in and out of this group than its 4- and 5-item counterparts.

Table 4.12: Predictive accuracy of the short form scales

Scale	Class	Sensitivity	Specificity	Balanced Accuracy
5- item form	5-7	0.93	0.97	0.95
	8-11	0.85	0.92	0.88
	12-15	0.8	0.96	0.88
	16-20	0.84	0.99	0.91
4- item form	4-6	0.94	0.94	0.94
	7-8	0.77	0.91	0.84
	9-12	0.72	0.96	0.84
	13-16	0.8	0.99	0.89
3- item form	3-4	0.9	0.97	0.93
	5-6	0.72	0.91	0.82
	7-9	0.83	0.91	0.87
	10-12	0.74	0.99	0.86

4.4 Regression overview

Regression is a statistical technique that reveals the association between an outcome variable (GESS, Contempt and Ostracism) and characteristics of interest, such as socio-demographic characteristics. The value of regression is that it reveals the association between each input variable (characteristic of interest) and the outcome variable, taking account of any relationship the input variable may share with another input variable and the outcome variable.³² It does not tell us about causality; rather it highlights relationships between two or more variable. Nevertheless, this makes it a powerful tool to understand the impact of differing variables on stigma.

Separate models were run for each of the outcome variables, i.e., GESS, Contempt and Ostracism. The results reported are from a standard Ordinary Least Squares (OLS) regression model, which assumes a continuous outcome variable. A further model was run for GESS using a more advanced specification of the model specified for a distribution of non-negative integers. The results of this generalised linear model were consistent with the standard OLS regression results, so the latter are reported given the technique is more familiar to a general audience.

The following variables were included in the model:

- Children in the household (reference = no, coefficient = yes)
- Gender (Reference = male, coefficient = female)
- Age (continuous)
- Work status (reference = no, coefficient = yes)
- Income (reference = below £50k per annum, coefficient = £50k+)
- Ethnicity (reference = non-white, coefficient = white)

³² Regression should not be confused with a Key Drivers Analysis (KDA). KDA attempts to rank explanatory variables in order of importance of their association with the outcome and does so using only the unique association between each explanatory variable and the outcome, i.e. shared variance is dispensed with. Regression maintains shared variance in the model and consequently, that shared variance is allocated to one of the variables in the regression results.

- Religion (reference = none, coefficient = yes)
- Long-term health/disability (reference = no, coefficient = yes)

4.5 GESS

The R² value shows that 28% of variability in GESS scores is explained by the demographic variables included in the model, which is very reasonable for this type of study and for the type of variables included in the regression model. The strongest predictors of GESS scores were age, followed by disability and gender, whilst religion and presence of children had similar effect on GESS scores (adjusted for other variables in the model). The direction and size of the relationships in the model suggests that older respondents and females are less likely to have higher GESS scores, whilst the opposite is true of people who are religious or suffer from disability and long-term ill health.

Table 4.13: GESS model results

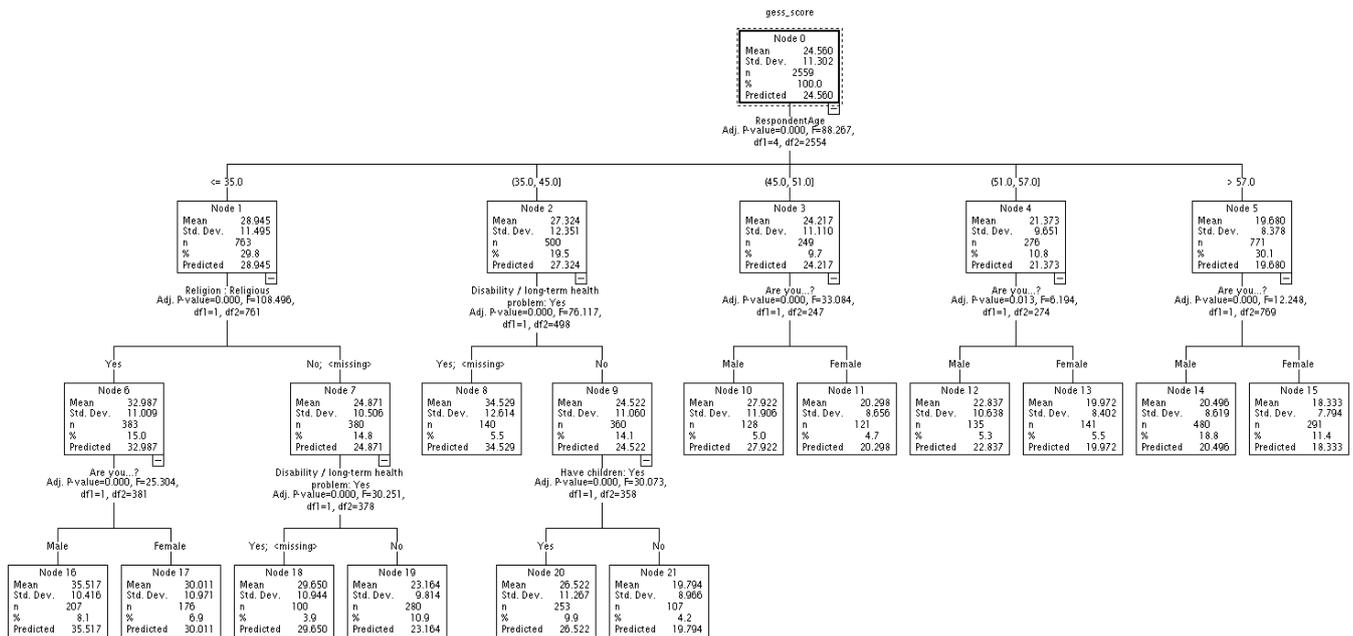
Characteristic	Standardized Coefficients	t	P	Zero-order	Part
Have children: Yes	0.173	8.525	<.001	0.338	0.151
Gender	-0.175	-9.621	<.001	-0.175	-0.171
Respondent Age	-0.262	-	<.001	-0.345	-0.215
Work status: Working	0.063	3.072	0.002	0.218	0.055
Income: £50K+	0.061	3.146	0.002	0.2	0.056
Ethnicity: White British	-0.051	-2.727	0.006	-0.128	-0.048
Religion: Religious	0.162	8.759	<.001	0.189	0.155
Disability / long-term health problem: Yes	0.198	10.854	<.001	0.207	0.193

Zero-order correlations (relationship between GESS score and individual demographics irrespective of impact of other demographics variables) demonstrate that although income (whether you earn over £50,000 p.a. or not) and work status have larger bi-variate co-efficient with GESS score than religion, their relationship with GESS scores is already explained by age (older groups are less likely to be in work) and by having children (those with children are three times more likely to earn more than £50,000).

Regression can offer us an understanding of the main predictors of GESS scores, but it does not determine the main segments of respondents most or least likely to have high GESS scores. Thus, CHAID was deployed to identify these groups. CHAID splits the sample into a series of subgroups that share similar characteristics called a “decision tree”.

As with regression results, the most discriminating demographic variable is age, followed by disability, gender and religious. A segment group or node with the highest GESS score is comprised of under-35 religious males (8.1% of the sample), followed by middle aged respondents with disability etc. On the

other hand, the segment groups least likely to have high GESS scores are elderly females, or middle-aged people who are not disabled and don't have children.



5 GPSS Technical Annex

5.1 Impact of reductive vs non-reductive language in the questionnaire

Two versions of the GPSS items were presented. The standard version referred to “problem gamblers” (PG) and the alternative version referenced “people experiencing harms from gambling” (PEHG).

As discussed, the term “problem gamblers” is considered reductive (in contrast to the non-reductive “people experiencing”) and therefore undermines GambleAware’s wider commitment to use appropriate language that help reduce stigma. GambleAware hopes to transition to the non-reductive language.

With this in mind, an experiment aimed to explore the impacts of varying the wording to identify:

- How, if at all, the change in wording affected how participants scored each item on the GPSS.
- Whether the change in wording affected how items on the scale correlated (e.g., a participant scoring high or low on one item was likely to score high or on another).

On average, the traditionally worded reductive version referencing “problem gamblers” tended to have a higher score than the alternative non-reductive version using “people experiencing harm from gambling”. However, even though 12 from the 13 items showed statistically significant differences between the two wording versions, the size of the difference was generally very small. For example, using Cohen’s d effect size as the measure of importance, only “Most people think x are liars” exceeded a value of 0.2, at

0.28. A standard interpretation of Cohen's d values is that 0.2 is small, 0.5 is medium and 0.8 is large.³³ Consequently, taken individually, the impact of the wording is generally small or non-existent for all but one GPSS item.

Overall, despite the statistical significance of most of the wording differences, our expectation is that these generally are too small to have a major impact on further analysis. There is some evidence of cumulation when summing across the items to calculate the summary scores. However, even here, the differences are around 0.5 of a scale point. We considered the potential for reducing the PG subsample scores but decided against this because adding a decrement of around 0.5 to the PG summary score would lead to inconsistency in the scoring and reduce the transparency of the results. However, Section 3.2 explores the structural integrity of the two scales using confirmatory factor analysis to check upon equivalence of the model between the two variant subsamples.

Table 5.1: GPSS item scores by question version

Question	Mean scores					
	"Problem Gambler"	"People experiencing harms from gambling"	Diff	P	P < 0.05	Effect size
Most people think {#GPSS_insert1} are liars	2.85	2.63	0.22	0.000	Y	0.28
Once they know a person is {#GPSS_insert4}, most people will take his or her opinion less seriously	2.81	2.74	0.08	0.001	Y	0.11
Most people think that {#GPSS_insert1} tend to be unreliable	3.00	2.89	0.11	0.000	Y	0.16
Most people think that {#GPSS_insert1} are unable to handle responsibility	2.94	2.86	0.08	0.000	Y	0.11
Most people think that {#GPSS_insert1} are lazy	2.52	2.45	0.07	0.004	Y	0.09

33

[https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3840331/#:~:text=Interpreting%20cohen's%20d&text=A%20commonly%20used%20interpretation%20is,rigidly%20\(Thompson%2C%202007\).](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3840331/#:~:text=Interpreting%20cohen's%20d&text=A%20commonly%20used%20interpretation%20is,rigidly%20(Thompson%2C%202007).)

Most people think that {#GPSS_insert1} are greedy	2.64	2.57	0.08	0.003	Y	0.10
Most people believe that {#GPSS_insert1} have no self-control	3.10	3.01	0.09	0.000	Y	0.12
Many people would be uncomfortable communicating with a {#GPSS_insert2}	2.79	2.79	0.01	0.749	N	0.01
Most people think less of a {#GPSS_insert2}	2.98	2.89	0.10	0.000	Y	0.14
Most people would not hire a {#GPSS_insert2} to take care of their children	3.07	3.02	0.05	0.045	Y	0.06
Most people would be suspicious of a person if they knew they were a {#GPSS_insert2}	3.03	2.95	0.08	0.000	Y	0.12
Most people would not want to enter into a committed relationship with someone they knew {#GPSS_insert3}	3.18	3.11	0.07	0.002	Y	0.10
Many people would avoid a person who {#GPSS_insert3}	2.93	2.84	0.09	0.000	Y	0.13
Summary Score: Contempt	16.76	16.13	0.63	.001	Y	0.20
Summary Score: Ostracism	21.09	20.61	0.48	.001	Y	0.13

Note: Effect size refers to Cohen's d.

5.2 Assessing structural integrity

Previous research provided evidence that the 13 items represented two subscales, i.e., Ostracism and Contempt. We explored this using a factor analysis of the 13 GPSS items across people with and without recent gambling experience, which did reveal the two dimensions. However, not all items cleanly measured either Ostracism or Contempt. Some items measured aspects of both GPSS subscales, i.e., there was cross-loading of items across the factors. In practice, it is likely that people showing a high degree of Ostracism are also likely to show a high level of Contempt (and vice-versa), i.e., the two dimensions are correlated. To check this assumption, the factor analysis was rerun allowing the two

factors to be correlated. However, this approach still did not resolve the issue of some items cross-loading across the two factors even when the correlation between the factors was accounted for.

Table 5.2: Factor analysis results for the combined GPSS items

	Question Item	Independent		Correlated	
		1	2	1	2
Original paper identified items as "contempt"	Most people think {#GPSS_insert1} are liars	0.34	0.62	0.17	0.59
	Once they know a person is {#GPSS_insert4}, most people will take his or her opinion less seriously	0.45	0.52	0.34	0.43
	Most people think that {#GPSS_insert1} tend to be unreliable	0.53	0.45	0.46	0.32
	Most people think that {#GPSS_insert1} are unable to handle responsibility	0.51	0.48	0.43	0.36
	Most people think that {#GPSS_insert1} are lazy	0.09	0.83	-0.19	0.93
	Most people think that {#GPSS_insert1} are greedy	0.14	0.80	-0.12	0.87
	Most people believe that {#GPSS_insert1} have no self-control	0.62	0.28	0.63	0.08
Original paper identified items as "ostracism"	Many people would be uncomfortable communicating with a {#GPSS_insert2}	0.53	0.43	0.47	0.29
	Most people think less of a {#GPSS_insert2}	0.65	0.33	0.65	0.13
	Most people would not hire a {#GPSS_insert2} to take care of their children	0.73	0.16	0.80	-0.10
	Most people would be suspicious of a person if they knew they were a {#GPSS_insert2}	0.72	0.28	0.75	0.05
	Most people would not want to enter into a committed relationship with someone they knew {#GPSS_insert3}	0.80	-0.02	0.95	-0.34
	Many people would avoid a person who {#GPSS_insert3}	0.66	0.35	0.65	0.15

Separate analysis of the two subsets of items for Ostracism and Contempt did show that the items, in each case, did load onto single factors. Both exploratory and confirmatory factor analysis was used separately for each subscale. The confirmatory models tested for both a common structure (a single underlying dimension) and common intercept (the same average score for each item) across the two wording variants. The results showed that the single factor model was appropriate in both cases with equivalent loadings of items on across the two wording variants. However, there was a significant difference in the intercept values, confirming the tendency for responses to be slightly higher among respondents viewing the "problem gambling" question word version than the version using the alternative people with gambling harms.

Table 5.3: Confirmatory factor analysis model significance tests

GPSS Construct	Model	Chi SQ	df	P
Contempt	Fixed coefficients	4.57	6	0.60

	Full equivalence	43.59	6	0.00
Ostracism	Fixed coefficients	5.52	5	0.36
	Full equivalence	19.61	5	0.00

Table 5.4: Confirmatory factor analysis loadings: Contempt

Question	“Problem Gambler”		“People experiencing harms from gambling”	
	Intercept	Coefficient	Intercept	Coefficient
Most people think {#GPSS_insert1} are liars	2.85	1.00	2.63	1.00
Once they know a person is {#GPSS_insert4}, most people will take his or her opinion less seriously	2.81	0.90	2.73	0.90
Most people think that {#GPSS_insert1} tend to be unreliable	3.00	0.97	2.89	0.97
Most people think that {#GPSS_insert1} are unable to handle responsibility	2.94	0.99	2.86	0.99
Most people think that {#GPSS_insert1} are lazy	2.52	0.95	2.45	0.95
Most people think that {#GPSS_insert1} are greedy	2.64	1.01	2.57	1.01
Most people believe that {#GPSS_insert1} have no self-control	3.10	0.84	3.02	0.84

Table 5.5: Confirmatory factor analysis loadings: Ostracism

Question	“Problem Gambler”		“People experiencing harms from gambling”	
	Intercept	Coefficient	Intercept	Coefficient
Many people would be uncomfortable communicating with a {#GPSS_insert2}	2.79	1.00	2.78	1.00
Most people think less of a {#GPSS_insert2}	2.98	0.98	2.89	0.98
Most people would not hire a {#GPSS_insert2} to take care of their children	3.07	1.05	3.02	1.05
Most people would be suspicious of a person if they knew they were a {#GPSS_insert}	3.03	1.08	2.95	1.08
Most people would not want to enter into a committed relationship with someone they knew {#GPSS_inert}	3.19	0.96	3.11	0.96
Many people would avoid a person who {#GPSS_insert3}	2.93	1.11	2.84	1.11

Despite the small differences in average scores across the two versions of the question wording the two question wording subsamples were kept as a single combined sample for further analysis with no attempt to adjust for these small differences, for reasons described above.

Consequently, Cronbach's alpha measures for the internal consistency of responses were 0.82 for Contempt and 0.86 for Ostracism. These are both highly respectable scores indicating a high degree of consistency between the items. The corresponding measure of Cronbach's alpha for all 13 GPSS items was 0.9. Overall, it was concluded that despite some fuzziness around the loading of the indicators on the factor analysis across the 13 GPSS questions, the existing measures of Ostracism and Contempt would provide a sound basis for further work.

Both the Ostracism and Contempt subscales largely tend to reproduce the total GPSS score distribution. The correlation between the three scales shows that Ostracism correlates $r = 0.92$ with the GPSS total, and the corresponding $r = 0.94$ for Contempt. In other words, there is a close correspondence between a person's level of perceived Ostracism and a person's level of perceived Contempt and the total GPSS score. However, Ostracism and Contempt correlate less highly with each other, $r = 0.71$. Consequently, it appears that while many people's scores on Ostracism and Contempt will be in accord, others may score higher on one of the scales and lower on the other. For that reason, we have chosen to explore the two subscales separately in the following analysis rather than considering the overall GPSS score. If both subscales were completely in accord, they would effectively be measuring the same single underlying single dimension and there would be no practical need for two separate scales.

5.3 Grouping the summary score

Latent Class Analysis (LCA) was used to identify the appropriate number of groups to create for the summary subscale scores. The LCA was initially carried out on a randomly drawn subsample of the dataset (the training dataset) and the model was used to predict class membership on the remaining subsample of the data (the validation dataset).

The distribution of the summary score by latent class membership was then used to identify corresponding cut-off points for grouping the summary score directly rather than using the latent classes.

Separate LCA models on the Ostracism and Contempt items suggested a three-class model was most appropriate, for each scale.³⁴ The average total Ostracism score increased across the latent classes from around 13 for the Low class to 18 for the Moderate Class and 22 for the High class. Similarly for Contempt, the average total scores increased from 15, 19 and 24, respectively from Low to High. This pattern shows that the model is distinguishing people by their total scores on the two subscales rather than using a different pattern of responses to create the classes, e.g., low scores on some items grouping with high scores on other items. Thus, the above findings showing that the items work well as a summary score is further supported by these findings. It is also apparent that the pattern of results for each item comprising the Ostracism scale increases in line with the total scores across the latent classes. Whilst there is some degree of difference in the individual item scores the range is comparatively small varying between 2 and 2.5/2.6 in the Low class (for the Training and Validation subsamples, respectively).

³⁴ The LCA models have been run without adjusting for any potential effect of question wording variant. Effects on individual items were generally considered to be small in magnitude and of interest only when cumulated as a total score.

5.3.1 Ostracism

Table 5.6: Average GPSS Ostracism score by latent class group

Question	Training			Validation		
	Low	Mode rate	High	Low	Mode rate	High
Ostracism total score	13.3	17.8	21.9	13.4	17.8	21.9
Many people would be uncomfortable communicating with a {#GPSS_insert2}	2.1	2.8	3.5	2.1	2.8	3.4
Most people think less of a {#GPSS_insert2}	2.2	2.9	3.6	2.2	2.9	3.6
Most people would not hire a {#GPSS_insert2} to take care of their children	2.3	3.0	3.7	2.4	3.0	3.7
Most people would be suspicious of a person if they knew they were a {#GPSS_insert}	2.1	3.0	3.8	2.2	3.0	3.8
Most people would not want to enter into a committed relationship with someone they knew {#GPSS_inert}	2.5	3.1	3.7	2.6	3.1	3.7
Many people would avoid a person who {#GPSS_insert3}	2.0	2.9	3.6	2.0	2.9	3.6

Based on the distribution of the summary scores (Table 5.7), Ostracism was grouped into 6-15, 16-20 and 21-24.

Table 5.7 shows the distribution of the Ostracism summary scores by the latent classes for the validation subsample. A key finding is the tendency for few people to score “2” or under (“slightly disagree”) across all individual items, i.e., less than a summary score of 15. In general, many people seem to at least strongly agree to many of the items (a summary score of 21), resulting in a large number of people falling into the intermediate class.

Table 5.7: Ostracism score by latent class (validation subsample)

Ostracism	Latent Class	Total
-----------	--------------	-------

	1	2	3	
6.00	0	0	12	12
7.00	0	0	5	5
8.00	0	0	7	7
9.00	0	0	4	4
10.00	0	0	6	6
11.00	0	0	12	12
12.00	0	0	83	83
13.00	0	0	46	46
14.00	0	1	97	98
15.00	0	26	114	140
16.00	0	146	41	187
17.00	0	217	6	223
18.00	0	453	0	453
19.00	17	163	0	180
20.00	81	102	0	183
21.00	114	0	0	114
22.00	83	0	0	83
23.00	75	0	0	75
24.00	94	0	0	94
Total	464	1108	433	2005

5.3.2 Contempt

There was little variation in average item scores among the Low class (ranging between 1.9 and 2.1), although this range increased a little for the Moderate and High classes, suggesting a little more variation between people's responses in the Moderate and High classes, albeit relatively small by degree.

Table 5.8: Average GPSS Contempt score by latent class group

Question	Training			Validation		
	Low	Moderate	High	Low	Moderate	High
Contempt total score	14.6	19.4	24.3	14.8	19.5	24.2
Most people think {#GPSS_insert1} are liars	2.0	2.7	3.5	2.0	2.7	3.5
Once they know a person is {#GPSS_insert4}, most people will take his or her opinion less seriously	2.1	2.8	3.4	2.1	2.8	3.4
Most people think that {#GPSS_insert1} tend to be unreliable	2.1	3.0	3.7	2.2	3.0	3.7
Most people think that {#GPSS_insert1} are unable to handle responsibility	2.1	2.9	3.6	2.2	2.9	3.6
Most people think that {#GPSS_insert1} are lazy	1.9	2.5	3.2	1.9	2.5	3.1
Most people think that {#GPSS_insert1} are greedy	2.0	2.6	3.3	2.0	2.6	3.3
Most people believe that {#GPSS_insert1} have no self-control	2.5	3.1	3.6	2.4	3.0	3.6

Identification of cut-off points for the Contempt summary score followed the same process as was used for Ostracism. Contempt was grouped into 6-16, 17-21 and 22-28 based on the distribution of scores across classes (Table 5.9).³⁵

As with the Ostracism score, a greater proportion of people tended to be grouped in a relatively restricted range of 17-21, giving little scope for splitting the Moderate group further.

Table 5.9: Contempt score by latent class (validation subsample)

Contempt	Latent Class			Total
	1	2	3	
7.00	0	0	11	11
8.00	0	0	6	6
9.00	0	0	3	3
10.00	0	0	7	7
11.00	0	0	11	11
12.00	0	0	7	7
13.00	0	0	26	26
14.00	0	0	84	84
15.00	0	0	82	82
16.00	0	3	121	124
17.00	0	106	63	169
18.00	0	198	9	207
19.00	0	248	1	249

³⁵ A cut-off point of 16 for the Low group was not straightforward and “17” could have been used had we based the decision on percentages rather than raw numbers.

20.00	0	229	0	229
21.00	11	297	0	308
22.00	86	49	0	135
23.00	94	0	0	94
24.00	84	0	0	84
25.00	70	0	0	70
26.00	37	0	0	37
27.00	21	0	0	21
28.00	41	0	0	41
Total	444	1130	431	2005

5.4 Correspondence between GPSS scales

Comparing the grouped Ostracism and Contempt scores shows that while the majority of people who are low, moderate or high in Contempt are also correspondingly low, moderate or high in Ostracism, substantial minorities are not (Table 5.10). In other words, a substantial minority of people have slightly mixed perceptions towards Ostracism and Contempt. Consequently, perceptions of these two aspects of stigma do not necessarily move in the same direction and should be considered separately.

Table 5.10: Correspondence between the Grouped Ostracism and Contempt scale scores

Contempt	Ostracism			N
	6-15	16-20	21-24	
7-16	59%	38%	3%	717
17-21	15%	76%	9%	2330
22-28	3%	44%	53%	958

5.5 Short-form development

The rationale for short-form versions for Ostracism and Contempt is less compelling than it is for the GESS, given the latter has 13 items, whereas the two former items have 6 and 7 items, respectively. Consequently, only 3-item versions of a short-form were considered for Ostracism and Contempt.

5.5.1 Ostracism

The following items were included in the short-form:

- Many people would be uncomfortable communicating with a {#GPSS_insert2}.
- Most people would not hire a {#GPSS_insert2} to take care of their children.
- Most people would not want to enter into a committed relationship with someone they knew {#GPSS_insert3}.

The distribution of the short-form score across the groups formed by the summary score across all items suggested short-form groupings of 3-7, 8-10 and 11-12, although it is apparent that there are many misclassifications. Overall, this short form grouping reproduced 86% of the long form classification.

Figure 5.11: Short form score by long form grouped score

Short-form	Ostracism group (long-form)			Total
	6-15	16-20	21-24	
3.00	34	0	0	34
4.00	19	0	0	19
5.00	37	0	0	37
6.00	236	8	0	244
7.00	263	67	0	330
8.00	192	494	0	686
9.00	24	1238	13	1275
10.00	3	508	120	631
11.00	0	146	288	434
12.00	0	4	311	315

Total	808	2465	732	4005
-------	-----	------	-----	------

Standard classification statistics are based on around two states: positive and negative. Consequently, we consider for each the predictive capacity of the short-form scale in predicting a single group of the long-form scale and repeat that process three times – once for each of the groups. Where a prediction from the short-form agrees with a classification in the long-form scale, it may be that it belongs to the class (a positive prediction) or it belongs outside the class (a negative prediction).

From the perspective of true and false positive and negative predictions, we can create various classification statistics that help us understand where a short-form scale predicts well and where it does not and what inaccuracies may arise in its use.

Sensitivity takes as its reference the total number of cases in the reference scale group and asks how many of these cases are accurately identified by the short form. The low scoring short-form group has a comparatively low sensitivity score of 0.73, whereas sensitivity improves for the moderate group (0.84) and again for the high group (0.82). The challenge for the low group can be seen from Table 5.11 with a cut-off score of “7”, there are 192 people in the long-form group with a low score who obtain a score of “8” on the short-form which, in conjunction with those also scoring 9-12 on the short-form, gives a comparatively large number who are misclassified in the long-form low scoring group.

Specificity takes as a base the number outside of the target group of interest and asks how many of these are accurately identified. This is lowest for the moderate group, reflecting their larger size compared to the other two groups. Conversely, specificity is comparatively high for the low and high groups.

Balanced accuracy takes the average of sensitivity and specificity and, for all groups, and shows the high scoring group perform reasonably well with a score of 0.89.

Overall, the challenge for the three-item scale is mostly apparent from Table 5.11, wherever the cut-off is drawn the restricted range of the short form score will result in substantial misclassification.

Table 5.11: Classification accuracy of the Ostracism short form

Group	Sensitivity	Specificity	Balanced Accuracy
Low	0.73	0.98	0.85
Moderate	0.91	0.77	0.84
High	0.82	0.95	0.89

5.5.2 Contempt

The following items were included in the short form:

- Most people believe that {#GPSS_insert1} have no self-control.

- Once they know a person is {#GPSS_insert4}, most people will take his or her opinion less seriously.
- Most people think that {#GPSS_insert1} are lazy.

The distribution of the short form score across the long-form groups suggested a grouping of the short-form by 3-6, 7-9 and 10-12. Overall, this short form grouping correctly reproduced 82% of the long form classification. However, Table 5.12 shows that those 246 people scoring 7 on the short-form will be classified as Low on the long-form and moderate on the short-form. This is a substantial misclassification but the misclassification error will not be resolved by recoding the short-form group, only shifted to a different location.

Table 5.12: Short form score by long form grouped score

Short-form	Contempt group (long-form)			Total
	7-16	17-21	18-28	
3.00	42	0	0	42
4.00	20	1	0	21
5.00	55	5	0	60
6.00	291	60	0	351
7.00	246	439	2	687
8.00	55	904	39	998
9.00	8	828	230	1066
10.00	0	86	339	425
11.00	0	6	191	197
12.00	0	1	157	158
Total	717	2330	958	4005

The Contempt short-form scale performs reasonably well in terms of sensitivity for the Moderate group, with specificity high for Low and High groups.

Table 5.13: Classification accuracy of the short form

Group	Sensitivity	Specificity	Balanced Accuracy
Low	0.57	0.98	0.77
Moderate	0.93	0.65	0.79
High	0.72	0.97	0.84

5.6 GPSS

The Contempt and Ostracism models showed very low R^2 values, indicating that socio-demographic characteristics were, at best, only weakly related to these scores. For Contempt, the R^2 value was 0.04, i.e., 4% of the variance was explained. The R^2 for Ostracism was close to zero at 0.01, indicating hardly any association with the socio-demographic variables. Nevertheless, despite these low measures, there was some indication that Contempt was higher among those with children, religious people and those with a disability or health problem. Similarly, religion and ill-health tended towards higher Ostracism scores and there was an increase in the Ostracism score with age.

Table 5.14: Contempt regression results

Characteristic	B	Std. Error	Beta	T	P
(Constant)	20.598	0.391		52.673	0
Have children: Yes	0.447	0.145	0.058	3.084	0.002
Are you male/female...?	-0.227	0.124	-0.031	-1.837	0.066
Respondent Age	-0.023	0.004	-0.106	-5.139	<.001
Work status: Working	-0.08	0.149	-0.01	-0.536	0.592
Income: £50K+	0.275	0.151	0.033	1.825	0.068
Ethnicity: White British	-0.294	0.181	-0.029	-1.628	0.104
Religion: Religious	0.526	0.126	0.073	4.179	<.001
Disability / long-term health problem: Yes	0.858	0.138	0.106	6.196	<.001

Table 5.15: Ostracism regression results

Characteristic	B	Std. Error	Beta	t	P
(Constant)	16.915	0.347		48.706	0
Have children: Yes	0.237	0.129	0.035	1.84	0.066
Are you... male/female?	0.156	0.11	0.025	1.425	0.154
Respondent Age	0.011	0.004	0.057	2.734	0.006
Work status: Working	-0.133	0.132	-0.02	-1.004	0.315

Income: £50K+	0.233	0.134	0.032	1.744	0.081
Ethnicity: White British	-0.213	0.161	-0.024	-1.324	0.186
Religion: Religious	0.361	0.112	0.057	3.231	0.001
Disability / long-term health problem: Yes	0.441	0.123	0.062	3.586	<.001

Appendix

The table below provide a breakdown of GESS and GPSS scores by socio-demographic groups. The shading represents comparatively higher (darker red) and lower (lighter red) GESS scores.

High and very high experience of stigma within GESS by demographic characteristic

	% Experience high stigma	% Experience very high stigma
Male	30%	19%
Females	23%	10%
18-24	45%	17%
25-34	30%	21%
35-44	27%	21%
45-54	18%	11%
55-64	18%	3%
65+	13%	2%
White	24%	15%
Mixed/multiple ethnic groups*	20%	33%
Asian/Asian British	41%	25%
Black/African/Caribbean/Black British	50%	20%
Other ethnic group*	42%	14%
Activities limited a lot due to health condition/disability	27%	42%
Activities limited a little due to health condition/disability	30%	17%
Activities not limited due to health condition/disability	26%	10%
No religion	24%	9%
Christian	26%	17%
Muslim	41%	37%
Hindu*	44%	20%
Jewish*	33%	33%
Other*	29%	27%
Heterosexual	26%	15%
Gay or lesbian	27%	7%
Bisexual	34%	22%
Other sexual orientation	44%	8%
Gross income <£13,499	32%	13%

Gross income £13,500-£24,999	28%	9%
Gross income £25,000-£29,999	27%	12%
Gross income £30,000-£39,999	26%	10%
Gross income £40,000-£49,999	28%	13%
Gross income £50,000-£74,999	25%	16%
Gross income £75,000+	25%	35%
Has no children	23%	7%
Has 1 child	33%	22%
Has 2 children	29%	29%
Has 3 children	32%	26%
North East England	27%	12%
North West England	29%	13%
Yorkshire and Humberside	27%	14%
West Midlands	33%	13%
East Midlands	27%	13%
East of England	27%	10%
South West England	21%	8%
South East England	25%	13%
Greater London	28%	29%
Wales	20%	17%
Scotland	24%	7%
Working	29%	18%
Non-working	22%	7%
PGSI 0	7%	1%
PGSI 1-2	20%	4%
PGSI 3-7	39%	9%
PGSI 8+	51%	39%

* caution, low base <100

For more information

3 Thomas More Square
London
E1W 1YW

t: +44 (0)20 3059 5000

www.ipsos.com/en-uk
<http://twitter.com/ipsosUK>

About Ipsos Public Affairs

Ipsos Public Affairs works closely with national governments, local public services and the not-for-profit sector. Its c.200 research staff focus on public service and policy issues. Each has expertise in a particular part of the public sector, ensuring we have a detailed understanding of specific sectors and policy challenges. Combined with our methods and communications expertise, this helps ensure that our research makes a difference for decision makers and communities.

