



## **Public Voice**

Technical note on panel survey 25

Alan Turing Institute / Ada Lovelace Institute

January 2023

# 1. Technical note

## 1.1 Objectives

Public Voice panel survey 25 comprised a single survey on the topic of artificial intelligence. It was commissioned jointly by the Alan Turing Institute and Ada Lovelace Institute. The target population was GB individuals aged 18+ and living in residential accommodation. The target respondent sample size was 4,000.

## 1.2 Sample and fieldwork design

### 1.2.1 The Public Voice panel

At the time of this survey (November 2022), the Public Voice panel comprised 25,620 members from across the UK. Most of these panel members were recruited via the 'ABOS' method in which (probabilistically) sampled individuals complete a 20-minute recruitment questionnaire either by web or on paper. Recruitment surveys were carried out in 2019, 2020 and 2021 and the respondent samples have been linked together via a weighting protocol to form a single panel.

It is worth noting that in 2019 Kantar Public carried out a (smaller sample) interview-based recruitment survey to help 'benchmark' the parallel ABOS recruitment survey. The interview survey sample was compared to the ABOS recruitment survey sample with respect to every question, whether demographic or non-demographic. This allowed a multidimensional selection model to be generated, operationalised through a set of 'calibration' weights applied to the ABOS recruitment survey sample. These calibration weights are used to inform sampling from the panel for particular surveys. Kantar Public expects that this approach limits the influence of recruitment method on the validity of the survey estimates.<sup>1</sup>

A full technical report for Public Voice is available separately.

### 1.2.2 Sample design

The sample for the survey was drawn from among the 24,673 respondents to the Public Voice recruitment surveys who were (i) resident in Great Britain based on the latest information available, (ii) aged 18+, (iii) had joined the Public Voice panel, and (iv) had not left or been removed from the panel.

The eligible subset of the Public Voice panel was implicitly stratified by (i) sex/age group, (ii) highest educational level, and (iii) region before a systematic random sample was drawn.

The sampling probabilities applied to each panel member *varied* in an effort to produce a maximally representative respondent sample. Broadly speaking, this sampling probability was proportionate to the product of (i) the panel member's recruitment survey weight, and (ii) one divided by the estimated probability of the panel member responding to *this* survey, conditional on their inclusion in the recruitment survey dataset. In technical terms, panel members were sampled with a probability proportionate to a size measure ('PPS').

The mean sampling probability among survey-eligible panel members was 0.44, the standard deviation was 0.32 and the range was 0.05-0.95. In total, 10,754 panel members were sampled for this survey.<sup>2</sup> One in five of these (2,172) was allocated to a reserve pool, with uniform allocation probability and this reserve pool was further subdivided into five batches of n=434-435. Systematic

---

<sup>1</sup> The Covid-19 pandemic affected recruitment because Kantar Public needed to carry out the second and third Public Voice recruitment surveys (in 2020 and 2021) without a benchmarking interview survey. The data from the original recruitment surveys in 2019 was utilised to estimate a *period effect* that could be layered across the original calibration model.

<sup>2</sup> The mean sampling probability among *sampled* panel members was much higher than for panel members in general (0.67 compared to 0.44). This difference is a reflection of the disproportionate sample design described in the text.

sampling methods were used for both steps to ensure that the main sample and each reserve batch was a statistical replicate of the whole sample.

In the event, the first of the five reserve batches was issued just before the end of fieldwork to ensure that the target respondent sample size (4,000) was achieved by the fieldwork deadline (21/12/2022).

Within the main sample, a subset of 500 (drawn from among those who could be surveyed by web) was allocated to a 'soft launch', issued to the field two days before the remaining sample. This allowed the research team to check for any questionnaire scripting errors that had not previously been identified. No errors were found and the remaining 8,082 panel members were issued to the field as scheduled.

In total, 9,017 panel members were issued to the field: 8,582 from the main sample pool (including 500 in the soft launch phase) and 435 from the reserve pool.

### 1.2.3 Fieldwork design

The sample was issued to the field in three stages: (i) a random subsample of 500 - taken from the main issue sample of 8,582 – was issued as a 'soft launch' 28/11/2022; (ii) the remainder of the main issue sample (n=8,082) was issued 30/11/2022; batch 1 of the reserve sample (n=435) was issued 19/12/2022.

All fieldwork was completed by 21/12/2022.

Both web surveying and telephone interviewing methods were used for this survey. Although web surveying is the default method, a targeted subset of web non-respondents was eligible for telephone interview.<sup>3</sup> Furthermore, another subset of 290 panel members were issued straight to telephone on the basis of their previously recorded preference and/or their offline status. Due to scheduling constraints, web surveying was the only method offered to the 435 panel members in the issued reserve sample.

The fieldwork process was as follows:

28/11/2022	Main sample 'soft launch' invitation sent (email/SMS)
30/11/2022	Main sample 'full launch' invitation sent (email/SMS; letter if neither); telephone interviewing begins with those in the 'straight to telephone' subset
5/12/2022	Main sample first e-reminder sent (email/SMS)
7-8/12/2022	Main sample letter reminder sent (mailed letter; landed no earlier than 12/12/2022); telephone interviewing begins with non-responders eligible to be offered this mode
12/12/2022	Main sample second e-reminder sent (email/SMS)
15/12/2022	Main sample third e-reminder sent (email/SMS)
19/12/2022	Reserve sample invitation sent (email/SMS)
21/12/2022	All fieldwork ends

All emails contained individualised survey hyperlinks, so no login details were required. Additional verification was based on panel member birthdate (including year). Where an email address was available, SMS text messages were used only as a supporting communication, sent to panel members who had not opened the email 24 hours after it had been sent. The first three of these text messages did not contain an embedded survey hyperlink. However, the last text message (sent 15/12/2022) *did* contain a survey hyperlink. Where an email address was not available, all text messages contained a survey hyperlink. The contact management system *sendinblue*<sup>4</sup> was used for all email and text message communications.

A letter was sent to all main sample non-responders 7-8/12/2022. The letter contained survey login details but not a printed individualised survey hyperlink.

---

<sup>3</sup> This targeted subset was pre-allocated after the sample was drawn and initially comprised 3,157 panel members in the main issue sample. A telephone number was available for c.2,000 other panel members but telephone interviewing resources were focused on those least likely to respond by web, given the scheduling constraints.

<sup>4</sup> <https://www.sendinblue.com/>

Those who completed the survey – and for whom an email or mobile telephone number was available - were sent a thank you email and/or text message together with a £10 e-voucher. Those for whom no email or mobile telephone number was available were sent a thank you letter, including a £10 shopping voucher card.

### 1.3 Questionnaire design

The following steps were taken to produce the questionnaire.

An **inception workshop** between the Research Team and client was held to establish and confirm the overall objectives and scope of the survey content. During this meeting concerns were raised that the proposed scope for the questionnaire would not fit with the 20-minute online interview length. A solution was agreed upon, to create two streams of questions within the questionnaire so that all questions required would be covered whilst keeping to the 20-minute online interview length. In addition, technology cases use's that were not vital to the study would also be dropped.

To aid this process we supplied a **questionnaire timing tool**. This tool made it possible to estimate questionnaire interview length as and when updates to the questionnaire design were considered. This informed decision-making regarding how many questions per case use and how many case uses to include. It also ensured that the length of interview for the two streams of the questionnaire (Groups A and B) would match.

Prior to drafting the questionnaire, we supplied a **questionnaire map** which focused on the key 'across' questions that would be repeated across the technology case uses. Once this was agreed upon, an initial draft of questionnaire was supplied to the client, starting an **iterative review** and redrafting process. Given the complexity of the questionnaire and the need to clearly define a range of complex and relatively unknown technology case uses, aspects of the drafting process (namely, the initial drafting of the case use descriptions) were supplied by the client.

Questions were drafted based on **key design principles**: they should work for online self-completion; they should work across all screen sizes; and they should be written in plain English. And a preliminary strategic review of these questions was held to ensure they meet the policy requirement.

**Cognitive testing** was performed in order explore how participants understand, mentally process, and respond to questions and so help to identify and address problems in this process.

A total of **12 cognitive interviews** were performed over Zoom. Since the start of Covid-19, we have successfully conducted cognitive interviews remotely via Zoom for clients such as BEIS, US State Department, Historic England and UCL and we used this approach here. To identify suitable participants to include in cognitive testing quotas (such as age, gender, disability, socio-economic grade and region) were agreed in advance and adhered to. Our researchers draw on established assessment techniques, including verbal probing, think-aloud and paraphrase protocols, and observation of non-verbal behaviours.

Analysis from the cognitive testing was incorporated into the questionnaire design and a final **questionnaire signed off** with the client. Thereafter, a telephone version of the questionnaire was created catering for the telephone interviews and a Public Voice start and end module were also added to the questionnaire. These modules ensured panellists correctly signed into the questionnaire as well as also sought updates of (i) where the panel member lives, (ii) working status, and (iii) contact details. For the most part, the respondents needed only to confirm details already supplied but any changes were collected here. A 'current status' data file is kept for all panel members, initially equal to the recruitment survey status but updated where applicable.

Finally, the client was given the opportunity to review the questionnaire once it had been scripted and prior to soft launch. Further amends were incorporated at this stage at the client's request before signing off again and launching the questionnaire.

### 1.4 Fieldwork performance

In total, 4,012 questionnaires were completed and passed a basic QC test. The QC test had two components: (i) the respondent had to have completed the last substantive question, and (ii) the complete questionnaire length had to be at least 40% of the 'adjusted median' among web survey

respondents. The ‘adjusted median’ was the median among web survey respondents but excluding the longest 10% (which were mostly multi-session responses). There were 130 web responses that passed QC test component (i) but not (ii). Module allocation pattern had a negligible influence on questionnaire length.

Of the 4,012 passing the basic QC test, 3,759 responded by web (94%) and 253 by telephone (6%).

The overall conversion rate (the number completing the survey and passing the QC test divided by the number issued for fieldwork) was 44%. If a design weight is computed that is equal to the inverse of the survey sampling probability, then the conversion rate was 45% (the difference is due to the fact that those expected to respond at a lower-than-average rate were sampled at a higher-than-average rate, and vice versa). The conversion rate for the main issue sample was very slightly lower than anticipated, hence the late decision to issue batch 1 of the reserve sample. The conversion rate for the main issue sample – with its completed fieldwork protocol – was 46% (47% weighted).

In total, 3,447 panel members in the main issue sample were eligible to be called by telephone, although many of these completed the survey by web before a call was made. 2,788 were called at least once by a telephone interviewer, although a large number of these (803) ended up completing the survey by web too. Of those called at least once, 43% finished with an ‘interim’ outcome meaning that, technically, more calls could have been made had fieldwork not closed on 21/12/2022. However, the median number of calls made to these panel members was 8 (mean = 7.6) and, overall, only 3% of those panel members called by telephone ended up with an interim outcome *and* were called fewer than four times.<sup>5</sup>

Table 1.1 shows the final disposition of all issued cases.

*Table 1.1: Final disposition of all issued cases*

	N	%
<b>Issued</b>	9,017	100%
<b>Web questionnaire completed and passed QC</b>	3759	42%
<b>Telephone questionnaire completed and passed QC</b>	253	3%
<b>No web completion, issued to telephone, non-interview final outcome</b>	531	6%
<b>No web completion, issued to telephone, no final outcome</b>	1,201	13%
<b>No web completion, not issued to telephone</b>	3,273	36%

## 1.5 Weighting

The respondent sample was weighted in three stages:

- 1) For every respondent, a *base weight* was calculated that was equal to his/her recruitment survey weight divided by the probability of being sampled for the survey (which varied substantially).
- 2) For every respondent, a *propensity score weight* was estimated, as a function of the recruitment survey variables. Technically, this propensity score weight was equal to the estimated odds of being present in the fully weighted recruitment survey dataset *rather than* the base-weighted respondent dataset when the latter dataset is added to the former (meaning that survey respondents are present in both datasets). To limit over-reliance on the model, the propensity score weight was limited to the inter-95%ile range, and the value of the

<sup>5</sup> The number of calls is actually the number of call-periods in which a call was made. Each day is divided into three call periods: morning, afternoon, and evening.

product of the base and propensity score weights was similarly trimmed.<sup>6</sup> This was used as weight (2).

- 3) Using weight (2) as a starting point, the respondent sample was calibrated to the weighted ONS *Labour Force Survey* of July to September 2022 (the latest available) with respect to sex\*age group, region, birth country, and highest educational level. The classic raking algorithm was used for this step.

Based on the standard set of 362 category-level proportions across 99 recruitment survey variables, the median difference between the weighted respondent dataset and the weighted recruitment survey dataset was only 0.5 percentage points. 95% of differences were  $\leq 2$  percentage points.

The overall weighting efficiency was 75%, equivalent to a design effect of 1.33 and an effective sample size of 3,008 ( $4,012 \times 75\%$ ).

Table 1.2 shows the calibration matrix that was used for the survey, derived from the ONS *Labour Force Survey* of July through September 2022 but with some adjustments to reflect minor differences between the LFS and Public Voice variables.

*Table 1.3: ONS Labour Force Survey population estimates, July through September 2022, GB adults aged 18+*

Variable	Category	% of population
Sex/age group	All	100.0
	Male 18-24	5.2
	Male 25-34	8.6
	Male 35-44	8.1
	Male 45-54	8.2
	Male 55-64	7.9
	Male 65-74	6.2
	Male 75+	4.8
	Female 18-24	5.0
	Female 25-34	8.4
	Female 35-44	8.2
	Female 45-54	8.4
	Female 55-64	8.2
	Female 65-74	6.7
	Female 75+	5.9
	*Other	0.2
Region	NE England	4.1
	NW England	11.2
	Yorkshire & The Humber	8.4
	E Midlands	7.4
	W Midlands	9.0
	E England	9.7
	London	13.8

<sup>6</sup> The respondent sample was divided into five equal-sized groups on the basis of survey sampling probability: weight (2) was trimmed to between 1/3 and 3 times the median weight for the relevant 'sampling probability' group. This approach was taken to ensure that the probability of a trimmed weight was not related to a panel member's sampling probability.

	SE England	14.1
	SW England	8.8
	Wales	4.9
	Scotland	8.6
<b>Highest education level</b>	Degree level qualifications, aged 18-69	31.6
	Lower qualifications, aged 18-69	45.6
	No qualifications, aged 18-69	5.6
	Aged 70+	17.2
<b>UK birth/citizenship status</b>	UK born	82.8
	Not UK born	17.2

\* Declaration of sex as 'identify differently' fixed at recruitment survey weighted level