Preface

The Customer Satisfaction Tracker is run by Critical Research on behalf of Ofcom. The objective of the survey is to track and understand the attitudes of residential consumers to the quality of service they receive for each specific telecommunications service or product they are decision maker for within the household. It focuses specifically on levels of satisfaction with aspects of their service across the four key communication markets (landline, mobile, fixed broadband and pay TV). The study is designed to report on satisfaction at an overall level per market and by provider within market, where a provider has a market share of at least 4%.

The data was collected through an initial face-to-face study, followed by an online study. The primary aim of the online study was to increase the number of interviews achieved with customers of communications providers (with a market share of at least 4%) where fewer than 100 interviews were achieved on the initial face- to-face study.

For the face-to-face study Critical Research interviewed a quota sample of 2,547 adults aged 16+. Interviews were carried out across 200 different sampling points in the UK. In 2019, in order to report on any differences by nation across the four key communication markets, oversampling was conducted in each of the devolved nations with a total of 1,528 interviews achieved in England, 335 interviews in Scotland and 342 interviews in each of Wales and Northern Ireland. This oversampling in the devolved nations was corrected for in the data weighting. All interviews were conducted between 3rd and the 31st January 2019.

The online study was conducted through a research panel with 831 adults aged 16+. These online interviews consisted of 502 ‘general’ interviews and 329 ‘boost’ interviews. The ‘boost’ interviews, as mentioned above, were conducted with customers of those providers where fewer than 100 interviews were achieved on the initial face to face study. Additional interviews were also conducted with TalkTalk broadband customers in order to achieve at least 400 interviews in total among this particular group. These additional interviews were then weighted accordingly in the final data set.

The purpose of conducting the 502 ‘general’ online interviews was to assess any differences between the results from the face to face interviews and the online boost interviews that could be attributable to the different data collection methodology; and to align the results across the studies. The approach taken to align the results is discussed further in the Weighting section below.

The online interviews were conducted between 22nd January and the 6th February 2019.

Across both the face to face and online studies questions were asked upfront to establish household ownership of landline, mobile phone, fixed broadband and pay TV. For each of these services the questionnaire established that the respondent is the decision maker for that service and also whether any of the households’ communications services are taken from the same provider.

Analysis was conducted by each total market i.e. fixed line, mobile, fixed broadband, pay TV and those purchasing services as a bundle. On completion of the interviews, weighting was applied for each market section (i.e. each of the above markets) of the survey using profiles from Ofcom’s Technology Tracker Survey.

Between 2010 and 2016 customer satisfaction levels across the communications markets were monitored via questions on Ofcom's annual Switching Tracker. One of the key considerations when designing the Customer Satisfaction Tracker was to ensure comparability between this study and the Switching Tracker in order to track how satisfaction levels have changed compared to 2016.

Details of the sampling frame, research methodology, and weighting procedures are outlined in the following pages. A note on statistical reliability is also included.

Sample Design - Random Location Quota Sampling

To ensure consistency with trend data from the Switching Tracker, the approach to sampling for the face to face study is the same as that for the Switching Tracker, using Output Areas (OAs) as the basic building block for sampling, then using quota control by four key variables (age, gender, SEG and working status) to control the sample interviewed within each sampling point.

First stage

The OAs in the UK were grouped into sampling units (SUs), which were then were stratified by region and rural/ urban:

- firstly, all the SUs were sorted by region,
- the SUs were then sorted within region by rural/urban.

This approach controls the urban/ rural fallout of the sample, so no further quota is imposed. The sample extracted was checked for close correspondence to the UK population on two key variables:

- Deprivation Index for the United Kingdom.
- Cable/ non-cabled area

Since region has been used as the first sorting variable, regional distribution of SUs will be more or less in proportion to the number of residential addresses in each region.

Second stage

The size of a SU is measured by the number of addresses it contains. The SUs were selected with a probability proportionate to size. This ensures that all households within an SU have an equal chance of being selected, regardless of the size of the SU in which a household is situated. The number of interviews per SU was 13.

Quotas

The following quotas were set (within each SU) to represent the population within that SU, which means the overall quotas across the UK will closely match the UK population. Quotas were set using 2011 Census data for Great Britain and Northern Ireland.

- Age (16-24, 25-44, 45-64,65-74,75+)
- Socio-economic grade (SEG)
- Gender
- Working status

Fieldwork

Interviewers were provided with specific addresses. The average SU contains around 130 households in England and Wales and 160 households in Scotland and Northern Ireland, thus affording tight control over the addresses the interviewers called at. All interviews were conducted in the home, using CAPI (Computer Assisted Personal Interviewing).

As detailed earlier, a separate online survey was conducted in 2018 using an online panel to reach internet users aged 16+. For the 'general' online interviews interlocking quotas were set to be broadly representative of UK internet users based on gender within age with further quotas set based on household socio-economic group and nation. For the 'boost' interviews no specific quotas were set. One of the criteria for inclusion was taking a service from one of the smaller communications providers where fewer than 100 interviews had been conducted in the initial face-to-face study. These providers were:

Landline: EE and Plusnet

Broadband: EE and Plusnet

Mobile: Giffgaff

Pay TV: TalkTalk

The other criteria for inclusion in the boost study was if decision makers for broadband had broadband from TalkTalk

Weighting

The approach to weighting the data in 2019 mirrored the approach for this study in 2018 when the weighting was first designed and is summarised below.

The data from the in-home study is combined with the data from the 'general' and 'boost' online surveys and is weighted. Information about the weighting process is outlined in more detail below.

The overall weighting plan is to replicate the weighting used on the Switching Tracker, which is to weight to targets by age, gender, SEG and services used (S4 in the current study), the one difference this year being that an extra tier of weighting was needed to correct the nation imbalance. However, the Switching Tracker is face to face only, and we need to ensure that the mixed method approach used here will not produce different results. In particular, it is known that online panels can fewer responses in the top satisfaction/ agreement category, and this could bias against suppliers where more interviews come from the online panel, so in particular those we need to boost.

We therefore conducted initial weighting analyses to understand the difference we get between the two methods and find a way to "correct" the online sample to give corresponding results to the face to face/CAPI sample. It is important to say that we are not judging the more accurate approach but trying to ensure that results from the 2018 survey are comparable with those from the 2016 Switching Tracker, so need to "correct" the overall 2018 data to the 2018 CAPI results.

The first step in this is to weight the CAPI interviews only to the above targets (including correcting nation), and then compare with the online panel results for the 'general', (as opposed to the boost samples). To do this, we filter the (weighted) CAPI sample on internet users (as all the panellists are), to produce a new set of target weights, then weight the online panel data to this same target. This makes the two samples identical by age, gender, SEG and services used.

Two immediate differences were apparent, in the profiles at M1 (type of mobile deal) and IN1 (Type of Broadband service). The online panel had higher proportions of prepay and faster broadband. Both these would be expected to affect the satisfaction ratings directly, and therefore we weighted in both cases to the CAPI profile.

This left differences between the two samples on a number of questions, including the overall ratings (across suppliers) for each of the four products, and also on the technology attitudes (QC1), attitudes to saving money (QC3) and depth of internet use (QC2). We wished to minimize the degree of weighting (having a view to effective sample size) and the key objective was to align the two samples on the satisfaction ratings. Thus we decided that weighting out these other differences (QC1, QC2 and QC3) was only justified if it produced a noticeable improvement in the alignment on the satisfaction rating scales. In no case was a clear improvement obtained. Therefore these three variables were not introduced into the weighting.

The ideal would have been to find a single weighting scheme that could be used across the data; for example, calculating corrective weights for each product, averaging these at the respondent level and then using that average weight. However, this made only a minor improvement to the comparisons. Therefore separate weighting schemes were used for each of the four products, and also for the "bundles" section. The approach in each case was to calculate a score per respondent for the two key ratings questions (overall satisfaction and value for money such as L1 and L2), scoring each as 9, 7, 5, 3 and 1 and adding the scores, then comparing the scores between the two samples. As an example, the distribution for Landline was as follows, where for information we also show the scores after the corrective weighting was applied:

Score for Landline	Face to face/ CAPI	Online initial	Online corrected
18	25.25%	23.03%	25.25%
16	18.20%	14.13%	18.08%
14	26.31%	25.39%	26.16%
12	13.46%	13.20%	13.54%
10	8.13%	8.88%	8.20%
0-8	8.66%	15.37%	8.76%

The final weighting scheme had three stages – two stages of pre-weighting, then a final stage of rim weighting (as outlined above and used on the Switching Tracker), to provide six sets of weights – one per product, one for bundles and one for the "common" questionnaire areas such as QC1-QC3, demographics etc.

The first stage of pre-weighting corrected the CAPI sample for the nation skews, and the online sample only by M1, IN1 and using the key ratings correction mentioned in the previous paragraph (for the first five set of weights). The second stage is to weight down those suppliers for which boost samples were necessary, using a simple weight calculated by the number of interviews in the CAPI and general online sample, divided by the overall total (including the boost interviews). This only applied to the first four sets of weights, and applies to all interviews involving boost samples.

The overall weighting, although relatively complex, was extremely efficient with the relationship between effective and unweighted samples being over 75% for all four services, lower than last year but driven by the skew by nation introduced this year.

Guide to Statistical Reliability

The variation between the sample results and the "true" values (the findings that would have been obtained if everyone had been interviewed) can be predicted from the sample sizes on which the results are based, and on the number of times that a particular answer is given. The confidence with which we can make this prediction is usually chosen to be 95%, that is, the chances are 95 in 100 that the "true" values will fall within a specified range. However, as the sample is weighted, we need to use the effective sample size[1] (ESS) rather than actual sample size to judge the accuracy of results. The following table compares ESS and actual samples for some of the main groups across all respondents.

	Actual	ESS
Total	3378	2579
URBANITY: URBAN	2891	2241
URBANITY: RURAL	478	332
GENDER: MALE	1646	1732
GENDER: FEMALE	1254	1326
AGE: 16-24	391	322
AGE: 25-44	1089	880
AGE: 45-64	1113	883
AGE: 65+	783	623
SEG: AB	784	603
SEG: C1	1058	805
SEG: C2	660	516
SEG:DE	876	672

[1] Effective Sample Size shown as Effective Weighted Sample in the data tables produced

The table below illustrates the required ranges for different sample sizes and percentage results at the “95% confidence interval”:

Approximate sampling tolerances applicable to percentages at or near these levels

Effective sample size	10% or 90% ±	20% or 80% ±	30% or 70% ±	40% or 60% ±	50% ±
2579 (Total)	1.2%	1.5%	1.8%	1.9%	1.9%
1732 (GENDER: MALE)	1.4%	1.9%	2.2%	2.3%	2.4%
805 (SEG - C1)	2.1%	2.8%	3.2%	3.4%	3.5%
332 (URBANITY: RURAL)	3.2%	4.3%	4.9%	5.3%	5.4%

For example, if 30% or 70% of a sample of 2,579 gives a particular answer, the chances are 95 in 100 that the “true” value will fall within the range of ± 1.8 percentage points from the sample results.

When results are compared between separate groups within a sample, different results may be obtained. The difference may be “real”, or it may occur by chance (because not everyone has been interviewed). To test if the difference is a real one – i.e. if it is “statistically significant” – we again have to know the size of the samples, the percentages giving a certain answer and the degree of confidence chosen. If we assume “95% confidence interval”, the difference between two sample results must be greater than the values given in the table below to be significant:

Differences required for significant at or near these percentages

Sample sizes being compared	10% or 90% ±	20% or 80% ±	30% or 70% ±	40% or 60% ±	50% ±
1732 vs. 1326 (Male vs. Female)	2.1%	2.9%	3.3%	3.5%	3.6%
603 vs. 805 (SEG AB vs. C1)	3.2%	4.2%	4.8%	5.2%	5.3%