SAMPLE SIZE GUIDANCE

TECHNICAL APPENDIX

This technical appendix accompanies the Sample Size Guidance for the Impact & Insight Toolkit (Toolkit) project. It describes the method used to derive the sample size guidance.

METHOD TO CALCULATE MARGIN OF ERROR

The Sample Size Guidance shows how many people to sample from a given audience size to achieve a given margin of error around estimates of mean average dimension scores. To produce this guidance, it was necessary to calculate the margin of error that would be achieved by taking different sizes of sample from a given total audience size.

The margin of error associated with different sample sizes was calculated using an aggregate dataset containing all the scores awarded to all the dimensions by all the public survey respondents in the Toolkit project to date. This dataset contains many thousands of dimension scores.

A set of example audience sizes was used. The example audience sizes were 50, 100, 250, 500 and 1000. For a given audience size, and for a given dimension, a 'dummy-audience' was created from the aggregate dataset. Samples of different sizes were taken from that dummy audience and the margin of error calculated each time. For each audience size the margin of error was calculated for every sample size possible for that audience. So for an audience size of 50, samples were taken of 1, 2, … , 49 people.

This process was replicated many times to find the 'typical' margin of error achieved by taking a given sample size from a given audience size for a given dimension. This type of procedure, where data is resampled over and over again, is known as a 'bootstrap method'.

The exact process to calculate the margin of error for a given audience size/sample size/dimension combination is:

1. Take a dummy-audience from the set of all responses for that dimension using random sampling **with replacement**. Find the mean of this dummy audience; this is the dummy population mean.
2. Take a sample from this dummy-audience using random sampling **without replacement** and find the mean of the sample.
3. Find the difference between the sample mean and the dummy population mean.
4. Repeat steps 2) and 3) a large number of times, recording the difference between the sample mean and the dummy population mean each time.
5. Convert the differences recorded above into margins of error by dividing each difference by the population mean and taking the absolute value.
6. Remove the 5% largest margins of error from list, reflecting a 95% confidence, and take the largest remaining margin of error.
7. Repeat Steps 1) through 6) a number of times (>50), generating a new dummy audience each time, and take the mean margin of error from all repeats.

The output of this process is a table of margins of error for each sample size for the selected audience size.

Table 1 – First 5 rows of margin of error table for audience size of 50

Sample Size	Margin of Error					
	Captivation	Challenge	Concept	Distinctiveness	Relevance	Rigour
1	53.7%	69.2%	54.4%	60.4%	67.6%	61.0%
2	37.6%	42.8%	34.7%	37.6%	41.7%	34.4%
3	28.6%	33.6%	25.3%	30.4%	33.3%	28.5%
4	24.6%	30.1%	22.6%	26.9%	30.1%	23.8%
5	21.0%	26.0%	19.5%	23.8%	27.0%	20.3%

Public surveys carried out as part of the Toolkit project normally contain all six of the dimension questions shown in Table 1. The recommended sample size required to achieve a given margin of error was taken as the sample size at which that margin of error was achieved for all six dimensions.

For example, Table 2 shows how, for an audience size of 50, the sample size required to achieve a margin of error of 5% is 40. This is because the Challenge dimension only achieves a margin of error of 5% at that threshold.

Table 2 - Example of how different dimensions offer different margins of error for the same sample size. Errors of 5% or over are highlighted.

Sample Size	Margin of Error					
	Captivation	Challenge	Concept	Distinctiveness	Relevance	Rigour
36	4.6%	5.8%	4.2%	5.0%	5.6%	4.5%
37	4.2%	5.1%	4.0%	4.5%	5.2%	4.0%
38	4.1%	5.2%	3.8%	4.5%	5.0%	4.3%
39	3.7%	5.0%	3.6%	4.2%	4.9%	3.8%
40	3.5%	4.6%	3.5%	4.1%	4.6%	3.4%
41	3.5%	4.1%	3.1%	3.9%	4.3%	3.3%

COUNTING WHAT COUNTS. CultureCounts