

Survey Errors

Survey questionnaires remain the most prevalent mode of data collection in marketing research, but has their high use resulted in diminished errors? Perhaps it is timely to revisit not the strengths – we know there are many – but the weaknesses that we should be cognisant of when analysing data collected using survey questionnaires.

The extent to which the survey data differs from the population it is collected to represent is described as ‘survey error’. This collective term describes the total survey error due to the following sources:

- Sampling error;
- Coverage error;
- Non-response error;
- Response (measurement) error; and
- Editing and processing errors.

Of these, sampling error is the only one that can be readily quantified, but all are important to consider. Sampling error is the loss of precision, statistically measured, that results from sampling rather than examining the entire population. It is the unavoidable trade-off

for the significant cost saving achieved through examining only part of the population, yet still being able to report results with acceptable precision and confidence.

Coverage error occurs when members of the relevant population are not given any chance of selection into the sample. Commonly this involves people who do not have a landline phone in telephone surveying. Since random digit dialling samples generally do not include mobile numbers, this problem affects those people who use only mobile phones.

Similarly, online panels are limited to the population with internet access. Another demographic bias that can affect online panels is the existence of ‘professional’ respondents who enrol to participate in surveys and can skew the population.

Non-response error can be categorised according to type: survey non-response and item non-response. Survey non-response is where a respondent who is eligible to complete the survey does not. This is primarily due to one of three

reasons: non-contactable, refusal or unable to participate. Item non-response is where one agrees to complete the questionnaire but leaves some items incomplete.

Whilst the error associated with non-response cannot be calculated, response rates can be derived and taken into account when setting sample size.

Response or measurement error is often the most problematic of all the survey errors. It occurs when the responses collected do not accurately reflect the true attributes of the respondent. This can result from problems with the instrument (poor wording, order or question response options, and ambiguity can be minimised by applying pilot and reliability testing), the interviewer (incorrect emphasis or recording of results can be minimised through training, supervision, implementation of CATI/CAPI) or the respondent (recall bias, social desirability bias, literacy and other problems can be minimised by applying pilot and reliability testing to questionnaire items).

Editing and processing errors are those that result from incorrect transcribing of collected responses into the analysis file.

Of the total survey error, sampling error seems to be the source that is most commonly raised by researchers. Understandably, the ability to statistically quantify this error lends itself to consideration. However, it is timely to remind ourselves not only of the existence of the other errors, but also of the importance of taking their influence into account.

“To err is human, to forgive is divine – but to include errors in your design is statistical.” – Kish, 1978.



[View this and other articles by Forethought online at *forethought.com.au/thinktank*](http://forethought.com.au/thinktank)