

Random samples in online panels

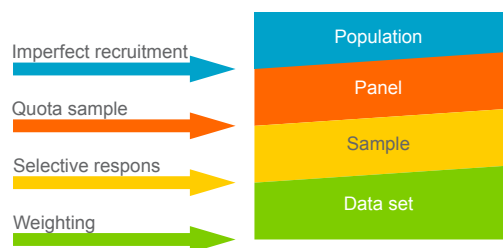
A new approach

Compared to quota sampling, including response inclination in the sampling process increases the validity of the unweighted net random sample. Compared to propensity sampling, response inclination results in a more efficient sample. This was the most important outcome from the study that Motivaction International conducted in cooperation with ProCresion and the University of Utrecht into the topic of random sampling in online research. The results of the study were presented on 29 October 2007 at the ESOMAR panel conference and the MIE 2007.*

Up to now

Random sampling is generally seen as the ideal sampling method.

However, this method is not useful when use with Internet panels, as an Internet panel is hardly ever an accurate reflection of the total population. Specific groups are either under- or over-represented in a panel. Therefore, random sampling from an Internet panel will result in a non-representative sample. For this reason, quota samples are often used in drawing Internet panels. Based on several characteristics—for example, age, sex, education—the gross sample is divided up among a number of sub-groups. Subsequently, a certain number of respondents or a quota is selected from each of the sub-groups. This results in a sample that is representative on the selected variables, but that is not necessarily representative on any other variable.



1: Quota sample

An advanced method used to obtain a more valid sample is propensity sample drawing. The major advantage of the propensity method over the quota sample method is that more—and different—variables may be incorporated into the sample drawing than simply sociodemographic variables.

With every sampling method, researchers have to contend with the occurrence of selective non-response. Initially, a sample may start out as representative of the target population. However, because respondents with specific characteristics choose not to participate in the study, the sample becomes a non-random sample of the target population. Due to this selective non-response, the quota sample, as well as the propensity sample, have to be weighted afterwards. However, weighting a sample reduces its efficiency and has unknown effects on validity. Therefore, it is important to draw a sample that requires a minimum of weighting.

One step further

If there is an online panel at your disposal, you know how often people have responded to invitations in the past to participate in a study. This 'historical response behavior' is the basis for the determination of response inclination.

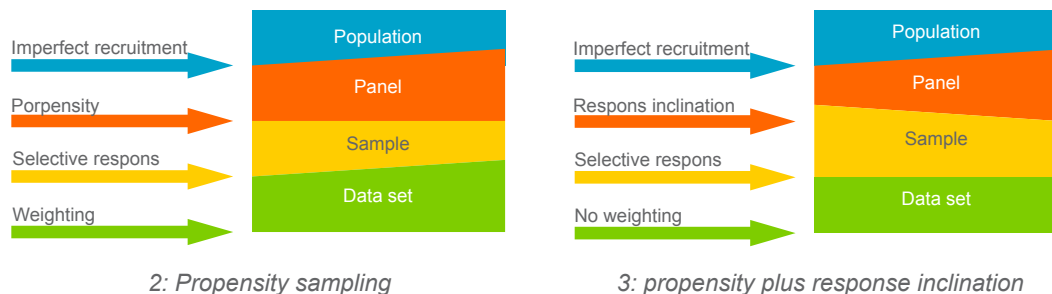
It is possible to ensure that the gross sample contains sufficient moderately-to-poorly responding panel members, which serves to limit the effect of the selective non-response. We know which panel members respond poorly to invitations to participate in a study. Integrating response inclination in the sampling process will lead to the invitation of more panelists that

motivaction

are more likely to have a poor response. This will result in a distortion of the distribution of the gross sample with respect to the population data. We expect a low response rate from this group, but since relatively poorly responding panel members have been invited to participate, the ultimate net sample becomes balanced in terms of its representativeness. This net sample hardly requires further weighting, which has a positive effect on the validity, representativeness and efficiency of the sample.

The research

Motivaction conducted an experiment with three types of sampling methods: quota sampling (1), propensity sampling (2) and propensity plus response inclination (3). Panel members received a questionnaire, including variables that were not used in the sampling methods. After evaluating the validity of the three samples, the distribution of these extra variables was compared with reference data on the Dutch population (Central Bureau of Statistics (CBS) data). Evaluation of validity was done on unweighted data.



Compared to the quota sample, the propensity plus response propensity sample was more valid and this did not occur at the expense of efficiency. In addition, after weighting, the propensity plus response propensity sample was more efficient than the propensity sample and had a similar degree of validity.

Among other methods, validity was measured using external reference data (CBS). The example below demonstrates that the propensity plus response propensity sample most closely matched the external reference data.

Statement: I am a religious person			
Percentage of affirmative responses ('yes')			
NL-population (CBS)	Quota	Propensity	Propensity+response propensity
41.5%	30.0%	32.9%	36.2%

Conclusion

Including response propensity to a propensity sampling increases the efficiency of the net sample compared to the traditional propensity sampling method. In addition the method of propensity plus response propensity has a higher degree of validity and, as a result, the quality of the net sample improves.

Martin Muller

Senior methodologist

* ESOMAR paper panel conference October 2007

Title: *Improving Panel Sampling: Embedding Propensity Scores & Response Behavior in Sampling Frames* by Lennart Huizing, Robert van Ossenbruggen, Martin Muller, Clasine van der Wal, Gerty J.L.M. Lensvelt-Mulders, Marianne Hubregtse.