Build Trust in Survey Responses
Sample Selection & Sample Size
Overview

One of the most commonly asked questions by clients during a market research study is, "What sample size do we need for results to be significant?" The answer is that meaningful results can be obtained with nearly any sample size; it's simply a question of what margin of error is acceptable. That margin of error is important in two circumstances:

1. Comparing results within a single survey
2. Comparing results between surveys conducted at different points in time

The structure of the sample itself is also important, particularly in B2B research where the population being sample can be much smaller than that for a consumer test. An incorrectly designed sample can be as disastrous to the validity of an analysis - if not more so - than an inadequate sample size.

So what can be done to mitigate these risks?

Selecting a Sample

The first, and perhaps most important, step in designing a market research study is to select a sample of respondents that represents the population of interest. This could be a representative sample of the national population, or could be a specific group within the population. Various sample providers have national, or even global, panels of willing research participants available to complete surveys using an incentive structure - typically a rewards systems based on points given for each completed survey.

Another option is to use internal lists, which is common for surveys of customers and employees. With the proliferation of online surveys, it is relatively easy to field a survey. Incentives are generally not required to get sufficient response among employees; usually a letter of encouragement from the CEO pointing out the benefits of employee participation will do. For

Sample size is one of many factors to consider when designing a survey for consumer or B2B research. Margin of error (MOE), or the range around a survey response in which we are reasonably confident that the "true" answer to a question would be if we asked that question to the entire population, is another. A term related to this is confidence interval (CI), which provides the confidence that the true response to a research question is within the margin of error, given the size of the sample being surveyed.
Build Trust in Survey Responses

customer surveys, an incentive (cash, gift card or charitable donation on the respondent's behalf) may be required. Even then, customer response can often be less than desired due to other priorities on their plates, so the sample may need to be enhanced with external lists from a B2B sample provider.

The next step is selecting the sample itself. When email lists are readily available, survey invitations can be sent to everyone on the employee/customer list with a valid email address as the cost per email is negligible. For surveys that require more resources, such as telephone or in-person surveys, there are a number of ways to select a representative sample, including:

- **Simple random sampling:** Select a random sample from a list using a random number generator, interviewing those who are selected.

- **Systematic sampling:** Interview every Nth person from a list (e.g. every 6th name). This is easier than simple random sampling since a random number generator is not needed, and is just as effective in creating a representative sample as long as there is no bias in the order of names used to generate the sample.

- **Stratified sampling:** Divide the population into subsets (e.g. geographical or demographics of interest) then select a sample from each group via either simple random or systematic sampling. This is a good method when there is interest in analyzing responses from distinct groups. For instance, healthcare clients are interested in gathering opinions from prospective patient both nationally and in the hospital's home market, so stratified sampling is ideal for achieving this objective.

During both fielding and analysis, it is important to avoid non-sampling errors, which include any error related to the research process. Field errors, which include factors such as interviewer bias or inattentive responses, are less prevalent today due to increased usage of online surveys. However, it is still important for survey designers to eliminate bias in the wording of questions as much as possible to get meaningful responses. Even when sample providers pre-screen lists for targeted surveys, screening questions in the questionnaire itself can help ensure that the proper respondents are being surveyed.

When practical, survey programming should specify that answers to questions are required to ensure that all respondents have the same interview experience. Asking the minimal number of questions to achieve research goals can help avoid survey fatigue that encourages inattentive or "straightline" responses, which greatly reduce data quality. Post-hoc analysis of individual
 responses or "trick" questions with obvious responses can be useful to catch inattentive response patterns.

One way to minimize non-sampling error can be to weight the sample as needed. Even with random selection and proper screening of respondents, there is still the potential for some groups to be over- or under-represented, creating a possible bias in results. This can be reduced either with setting quotas for subgroups when selecting the sample (see stratified sampling, above) or weighting responses during analysis to ensure proper representation. To perform post-hoc weighting, population data from the U.S. Census or similar sources can be applied to individual responses, essentially increasing the influence of under-represented groups and decreasing that of over-represented groups so that the average response more accurately reflects the population of interest.

**Sample Size**

Once the proper sample frame is in place, then it is time to address the next most important question: How many respondents should be in our sample? A simple answer is that the larger the sample, the more confidence we can have in responses. However, any research study using an external sample source will have cost constraints, as panel providers charge based on number of completed responses. In the case of specialized samples (physicians in a limited geographical area, for instance) panel providers only have so many potential panelists to "lease" for a survey so eventually you will hit a wall in terms of available respondents.

Given these constraints, we can still have confidence in our results. A properly selected sample with 30 or more respondents is typically treated as a normal distribution statistically, so t-tests and other measures of significance can be used to determine the meaningfulness of survey responses. Margin of error is very wide with small samples, so a sample of 30 can give meaningful results, but has an extremely wide margin of error so finding meaningful results is much more difficult than with a larger sample.

**Margin of error** (MOE) refers to the range around a survey response in which we are reasonably confident that the "true" answer to a question would be if we asked that question to the entire population. For example, a new product test of 1,000 decision-makers reports that Product A is preferred over Product B by 4% (52% vs. 48%), with a MOE of ±3 percentage points. We can be reasonably confident that preference for Product A is real, because its lead over Product B is more than the margin of error.
Build Trust in Survey Responses

Margin of error decreases as sample size increases. In the example above, a sample of 1,000 would give us a MOE of 3 percentage points. If we increased the sample to 1,500 the MOE would decrease to 2.5 percentage points, while lowering the sample to 250 would increase MOE to a little over 6 percentage points. If we saw the same results with a sample of 250, we could not be confident that Product A’s 4% preference lead is significant, and would declare the results of the survey to be a statistical tie.

The above numbers assume a confidence interval of 95%. Confidence interval (CI) refers to the percentage of time that we believe the "true" answer to a survey question falls within the margin of error. In the example below, we tested three TV advertising concepts with a sample of 800 respondents and 40% preferred Concept A, while 37% preferred Concept B. At a 95% CI, the MOE is ±3.5 percentage points, meaning we cannot be confident that Concept A is more preferred than Concept B using this CI. If we used a 90% CI instead, the MOE is ±2.9 percentage points so we could report a significant difference between Concepts A and B, though we are only 90% sure the difference is real. The level of confidence chosen depends on the level of comfort that an organization needs in showing true statistical differences in results, balanced with cost and other constraints in sample selection.
Build Trust in Survey Responses

The examples mentioned so far use proportions to demonstrate the concepts around sample selections, but these principles also apply to testing of means as well, in cases where means rather than proportions are reported in market research results. While sample size and MOE for proportions can be calculated using various online tools, making MOE for means requires standard deviation, a metric which is unknown until the study is performed and available in the output of statistical software when testing means. Assumptions can be made regarding standard deviation prior to fielding a study, but it is generally simpler to use proportions-based calculations to determine the sample size needed.

When calculating sample size, it is important to take into consideration which sub-samples may need to be compared in the analysis. Cost constraints may require accepting larger margins of error when comparing demographic groups or segments within the sample, but a rule of thumb is to plan for collecting samples of at least 100-150 respondents for sub-group that has a high level of interest within an organization.

Validation of Results

Statistical significance testing is the most important metric to consider when evaluating the results of a properly designed survey. However, there are a few other checks that should be made to ensure that internal and external clients are comfortable with survey results. If similar studies have been conducted, the researcher should be prepared to explain any meaningful differences between results of the current and past waves of the study. This is especially important for tracking studies, which are designed to measure attitudes and behaviors over time.

If results show any inconsistencies with past studies, a demographic comparison can be used to determine whether changes over time are meaningful, or simply a function of demographic variation among respondents over time. If demographics among the older and newer sets of respondents have changed, check these trends against U.S. Census or similar sources to see if this is a true demographic shift. If demographic differences are simply due to variation in samples, consider weighting responses of both data sets to Census norms to eliminate any bias.

Sometimes it is necessary to make a significant change to the methodology of a tracking study. Examples include changing research vendors, switching from telephone to online surveys, and making a major change to the number and types of questions. In these cases, we strongly recommend fielding the old and new studies in parallel until enough time has passed to either verify that results do not drastically change, or to fully understand how and why results change. If results change greatly between tracking studies, parallel testing should remain in place until prior
Build Trust in Survey Responses

Year comparisons are available in the new study to avoid disruptions to KPI measurements and business planning.

Taking Action

Both the quality of the sample and the number of people surveyed are important to ensuring the validity of results, which will help researchers present meaningful conclusions and recommendations to clients with confidence. Keeping these guidelines in mind will help ensure that the organization receives the most accurate information to address the business objectives.
About Endeavor Management

Endeavor Management is an international management consulting firm that collaboratively works with its clients to achieve greater value from their transformational business initiatives. Endeavor serves as a catalyst by providing pragmatic methodologies and industry expertise in Transformational Strategies, Operational Excellence, Organizational Effectiveness and Transformational Leadership.

Our clients include those responsible for:
  • Business Strategy
  • Marketing and Brand Strategy
  • Operations
  • Technology Deployment
  • Strategic Human Capital
  • Corporate Finance

The firm's 40 year heritage has produced a substantial portfolio of proven methodologies, deep operational insight and broad industry experience. This experience allows our team to quickly understand the dynamics of client companies and markets. Endeavor's clients span the globe and are typically leaders in their industries.

Gelb Consulting Group, a wholly owned subsidiary, monitors organizational performance and designs winning marketing strategies. Gelb helps organizations focus their marketing initiatives by fully understanding customer needs through proven strategic frameworks to guide marketing strategies, build trusted brands, deliver exceptional experiences and launch new products.

Our website:
www.endeavormgmt.com