Survey Research Terms

Survey Sample

When a survey is being implemented, the net "sample" is all respondents that get recorded into a data base. A questionnaire may get returned with no questions answered. In that case, it might be recorded as part of the returns, but it is not part of the net sample. The net sample is virtually always a subset of the population being surveyed. Very rarely is the entire population included in the results of a study.

Random Sample and Representative Sample

A random sample would be realized if every individual in the population would have an equal opportunity to respond and would be equally likely to respond. Equally likely to respond is challenging to determine as a number of variables will affect this such as reading levels and language as well as timing of the random sampling to name a few.

Since the bar to defend a sample as truly being random is so high, researchers rarely use that term but instead strive for the next best thing which is a representative sample. A representative sample is one to which RCS has applied some testing and feels is close to representative of the entire population on a variety of characteristics. RCS goes to great lengths to ensure that it can refer to its final survey samples as representative of the entire population, or comment on why and how it may over or under - represent certain subsets of it.

Statistical Reliability

Technically, the term statistical reliability can only be used if the sample is truly a random sample. However, most researchers make a reference to such statistical reliability even if the net sample cannot be defended as truly random. Reliability can be thought of as the likelihood of repeatability. A reliable survey sample is one that RCS can assert can be reliably replicated and therefore, from a research perspective, it can be relied upon it to give consistent answers that closely reflect the reality of what is happening in the overall population surveyed. If a survey project is repeated every five years and the answers are different, it can be assumed that the difference is "real" and not simply due to sampling error, or problems with the survey methodology. The degree of reliability is measured using a Confidence Interval.

Confidence Interval

A confident interval is a statistical calculation of how reliable a sample is deemed to be. The interval is expressed in two parts. A loose "industry standard" level of confidence that is most often sought is to be 95% confident that the results are with plus or minus 5% of perfectly reliable. That means that if a survey was repeated 20 times, in 95% of those times (19 out of 20) the answers would be within 5% of the answers in the initial survey.

As an example, if a survey is designed with a methodology to reach the industry standard threshold of reliability, and once the survey is completed it finds that 50% of respondents said that their household used a swimming pool within the past year, it could be concluded that somewhere between 45% and 55% of households actually used the pool. If the same survey conducted five years later found that 60% said that their household had used that pool within the past year, it can be concluded with 95% confidence that the increased proportion of household use is "real" because the increase of 10% (from 50% to 60%) is outside the margin of sampling error of plus or minus 5%. If, however, the increase over the five-year period between surveys were only 3%, it could not be concluded with any level of confidence that it is a "statistically significant" difference in the result over that period. In fact, one could not rely on the small difference as actually representing an increase in proportionate use.

Sample Size

Statistical reliability is always a function of sample size. If a sample is truly random or at least representative, the larger the sample the more reliable it is and the higher the confidence in its result. There are confidence interval tables that show the resulting level of confidence for all sample sizes. Those tables show that there is a "law of diminishing returns". If there are more than 25,000 adults in a population and the sample size is at least 1000, one could triple the sample size and not increase the resulting level of confidence significantly. Therefore, there is no need to spend more money trying to get 3000 returns once a study has realized the first 1000 survey returns. It may seem counterintuitive, but if one wants to know what thirty-five million Canadians are thinking or doing, once you have a random sample of 1000 of them, you don't need to increase your sample size to get a more reliable picture of what you want.

Validity

Validity is completely separate from reliability. Validity has to do with how information is collected. If survey questions include some inherent bias, the answers are not likely valid. Bias can be inserted into questionnaire development in a number of ways. If a researcher is trying to ascertain a respondent's physical address there is very little chance of introducing bias into the question. However, as an example, if there were a fee structure for use of a facility that had a higher fee for a non-resident, that could introduce an incentive for a respondent not to be truthful in their answer and that could render the results somewhat less valid.

Mean/Median/Standard Deviation

In survey research, if one is testing an opinion or a behaviour or characteristic of the respondent that has a range of answers along a continuum, a researcher can calculate a **mean** or average answer or a **median** which is the mid point in the range of all stated answers and subsequent frequency distribution of answers can be created which is a graph showing where on the continuum most answers fall. In these cases a Standard Deviation of answers can be calculated which illustrates how much variability there is in the answers. That is, do most answers fall close to the mean (a small standard deviation) or are answers spread over a wider range (a larger standard deviation).

For example, if one is asking about household income, there will be a mean along that series of answers which shows what the average household income is within the population surveyed. There will also be a median above which and below which the same number of households fall. And, a standard deviation can be calculated which shows how much "spread" there is on this single characteristic.

In the RDN usage study a respondent either uses a facility or doesn't. So, there is no continuum of answers, no mean, no median and no standard deviation of answers to calculate.

Mixing Data and Sample Types

In some cases, such as the RDN usage survey, research methodology can mix sources of data which complicates the mathematics of calculating confidence levels about where facility users live. One can have many sources of data with varying levels of reliability as follows:

- All program registrants over a year – which represents a perfect 100% population of all program registrants (rather than a sample of them) of all programs at a pool and therefore is 100% reliable. There is no level of confidence interval. We are 100% confident of this data, or as close to it as we can be.
- A complete data base of all drop-in users of a pool that paid via a purchased membership card that gets scanned for each drop-in to a public swim – as above, we have 100% confidence that we have a perfect reflection of all such drop-in users that paid using a membership card that requires a person to apply for the card with an address.
- A representative sample of 800 drop-in pool users drawn at different times of the week and in different months that has a calculated confidence interval of 95% plus or minus 4%.
- A list of all addresses of all members of most of the main pool user groups with a confidence interval of 95% plus or minus 5%

To get an overall confidence level in such a case, RCS combines all the levels of confidence in a way that reflects how much use is roughly associated with each source of data. In the above case, that confidence level would be *"better"* than the industry standard.

Summary of Methodology and Reporting of Usage Surveys

RCS is able to draw a number of conclusions from its work on the 2010 and 2015 usage survey.

1. The methodology used is sufficiently valid and reliable to be used to apportion net costs of operation for pools, arenas, and sports fields. While no data is perfect, the consultants assert that the information available and its analysis generate results which are more reliable and valid than industry standard levels of confidence. Industry standard level of confidence in survey data is plus or minus 5% nineteen times out of twenty. For this study, the combination of data sources with different levels of reliability are complicated to combine into a cohesive confidence level. However, the overall result is almost certainly within 2% nineteen times out of twenty.
2. This means that if the methodology were repeated consistently, use by area of residency would have to shift by more than 2% for it to be reliably picked up (nineteen times out of twenty) by the process.
3. The level of reliability is improving over time rendering results which are more reliable.

Page 3 from PERC 2015 Usage Report Re: Drop In Survey

There is no reason to indicate that the survey periods in February and May/June of 2015 were atypical of users or uses during other months of that year. There is also no reason to assume that the year 2015 is atypical of recent years. Therefore, the consultants believe that this methodology, which solicits residency from a large sample of facility users from each pool, is quite valid and reliability represents all drop-in users of each pool with an accuracy of about +/- 4% nineteen times out of twenty.