

Weighing your Options on Web Surveying? **Look, then Leap**

Doug Grisaffe, Ph.D.
Chief Research Methodologist,
Walker Information

Making decisions about collecting data on the web? The intent of this Walker Information white paper is to help you consider some of the issues regarding web-based surveys, based on studies of the methodology to-date. It will also offer recommendations for those trying to make decisions about web surveys versus more traditional modes of data collection such as mail and telephone.

Researchers should be wary in using blanket statements about the best data collection methodology, because it always depends on things like:

- Objectives of the research
- Population of interest
- Availability of sample
- Budget
- Timing requirements
- Subject Matter

If you want to survey voting preferences among people older than 80, a web-based survey is probably not appropriate. If you need top line information three days from now, a mail survey won't work. The best research always requires initial consideration of information needs and design factors. Only then can we decide if a web-based survey makes sense for a particular situation.

Enthusiasm for web-based research based on reality

Our first consideration is always a set of design factors. That said, Walker Information believes the research data to-date support web-based surveys as a viable option in many current research contexts. Some primary advantages in using the web include speed, cost, and ease of data collection. And, based on the best scientific research to date comparing electronic modes of data collection (e.g., web, e-mail) to more traditional modes, the evidence bolsters the case for web-based surveys.

Results indicate comparability of web and other methods with respect to response rates, data quality/missing values, average response values, and structural relationships among responses. There are some caveats to these conclusions, but they reflect the findings of the most well designed published tests.

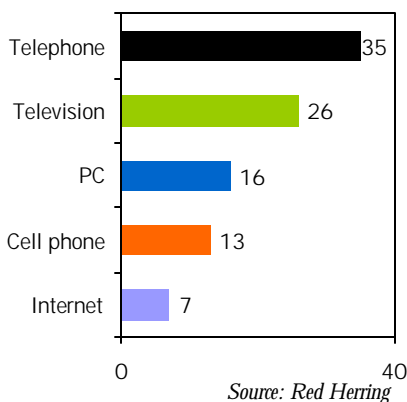
The big issue: representativeness

No matter how compelling the advantages of online surveys, there is one issue to consider closely before racing headlong to the web: representativeness.

We usually represent a population through sampling. After drawing a sample at random from a larger population, we administer our survey. Then, we project the survey results to the entire population, making an inference based on the sample. Certain principles from probability science work to our advantage in allowing us to make that leap. While there is some margin of error in our projections, when sampling is done correctly, the margin of error tends to be random and small.

Take to Tech:

Number of years for 1/3 of US households to adopt new technology



Representativeness may matter

For example, assume that there are age-related differences in breakfast cereal preferences. Perhaps younger adults favor sweetness, while older consumers favor fiber content. A researcher who does an online survey to get product development information for a new cereal offering may conclude that sweet is the way to go. The problem? He or she has based that conclusion on a biased sample. Since online users tend to be younger than the population at large, the preference for sweetness observed in the data does not accurately represent the full population. Instead, it only captures the "biased" preference of a special subset of the younger online population. The findings, and therefore new product decisions, don't hit the mark in the context of the entire market. In fact, given the vast numbers of retiring and aging senior adults in the US, major market opportunities would be missed through conclusions drawn blindly from the online methodology.

PROS*

- Fast/ compressed cycle
- Cheap/ elimination of traditional costs
- Visual and virtual stimuli
- No or low cost after setup
- No limit on number of completes per day
- Increased open-end participation
- Less intrusive than phone
- Good for certain hard-to-reach segments
- Participation at respondent's leisure
- Avoids threats to CATI (high resistance, answering machines, caller ID)
- No social acceptability bias (self-administered)
- Easily extends across national boundaries
- World wide consistency (no interviewer bias)
- Sensitive/embarrassing content easier
- Internet runs 24hrs /365 days

*Taken from published literature to-date

An online survey might not represent the population of interest as a whole

Online research faces a problem if some people in a population of interest have access to and use the Internet, while others do not. A survey done online will be representative only of those who are online, not of the population as a whole. A serious threat to validity is posed any time online and non-online people differ in meaningful ways. Therefore, a researcher considering web-based surveys should carefully evaluate the issue of representativeness.

At the time of this writing, online users are not representative of the US population as a whole (i.e. they tend to be younger, have higher incomes, and be male.) As long as this is the case, surveys collected online will not be truly representative. Representativeness, then, may be the greatest threat to validity in the online survey debate.

Massive online panels don't fully solve the problem

Some may be tempted to say "Forget it - web data is so fast and cheap it doesn't matter." But there may be important differences between individuals who are and are not online. Massive online panels, weighting of online results, or other potential approaches cannot fully solve the problem. An online panel is still just that - online. Weighting might help, but logically, it can only do so much. No amount of weighting can turn an "online person" into a non-online person.

Good News? The online gap is narrowing

The online population is increasingly mainstream as rising numbers of people (at least in the US) have access to the Internet. More people are coming online as access increases through cheaper PCs, Web TV, access through schools, PCs in businesses, and telecommuting setups in homes.

Also, there are many relevant research contexts in today's markets where representativeness is ensured due to the nature of the population of interest. For example, in the high-tech sector, many research populations are fully represented online. Web-based surveys in these sectors don't suffer the same threats regarding representativeness.

Finally, there are some research contexts where representativeness is not necessary. While that sounds like research heresy, consider the case of an organization facing breaking news on a managerial ethical lapse. Some quick online qualitative data might provide an excellent initial flavor of the variety of public reactions. There is no scientific probabilistic inference to be done here. It is simply a matter of getting some quick, early information on an array of public opinions.

Problems With Comparative Studies to Date

A number of published works exist in the literature covering online surveys. Many are conceptual discussions of the pros and cons of doing research online. Some studies have actual data to compare a variety of aspects of online vs. non-online surveys. Unfortunately, many of these data-driven studies are not fair tests. The empirical studies on this topic fall into two types: those that try to ensure equality/similarity of groups before collecting data, and those that do not. Let's consider the latter first.

Confounded Design: using test groups that are different from the start

If we start with groups that are inherently different at the outset, use a web method with one group and a phone method with the other, and then find some differences when survey results are compared across groups, what will be the true source of differences? Will it be that the web-based methodology made the results different? Or, will it be the initial group differences? You can't be sure. This is actually referred to as a "confounded" research design.

The two groups were different from the start. The results at the end are different. Is that attributable to the method of administration effect or the group difference effect? It is equally bad when the design just described produces no differences in results, and a conclusion is drawn that the mode of administration makes no difference. The mode of administration could be producing a 5 point upward bias for phone, whereas the starting group differences could create a 5 point downward bias due to group characteristics. These effects would then cancel out, making web and phone look equal, when really there could be a real bias that is missed due to the “confounded” design.

Good Design begins with homogenous test groups

The only true and fair test of the method effect comes from studies where some attempt has been made to make the groups equal before administering different survey modalities. Then, if we see differences or no differences in the results, we can be confident in attributing this solely to the method of data collection, (not group differences). Random assignment to conditions is the tried-and-true experimental means of ensuring group similarity from the start. Separate representative random samples from the same population, or, elaborate matching on characteristics (matched samples) are two other reasonable scientific approaches to starting similarity among conditions.

Bottom line – if you look to published research (or conduct your own) for answers regarding the comparability of web and other methods of data collection, the only fair test is one that exerts control in ensuring starting similarity/equality among the test conditions. If this is not done, the groups started out different, and it should be no surprise if they end up different and produce dissimilar survey results.

As noted earlier, some published studies have used scientific means of ensuring starting similarity among groups. In general these show similarity in response rates, data quality/missing values, average response values, and structural relationships among responses. Significant differences are seen in the areas of cost and speed – online surveys consistently show superiority over traditional methods.

CONS*

- Not accessible to all segments
- Non-representative “samples”
- Perceived security
- Duplication possibilities if no unique URLs/passwords
- No census lists, no RDD-like capability, no central registry of users
- Web-posted surveys have no control over respondent qualifications
- Some e-mail lists have false addresses
- Authentication difficult – false identities
- Interpretation as SPAM
- E-mail surveys not anonymous
- Privacy concerns may cause exit or false information
- “Elites” may prefer personal attention of live/in-person
- May be harder to do incentives
- Incentives may encourage duplication under different addresses
- E-pre-notification may lack credibility cues: logos, letterhead, signatures

*Taken from published literature to-date

Some Practical Recommendations

After determining that web surveying is a viable option, a researcher should decide whether it is reasonable to assume that there will be no effect on the scores due solely to mode of administration. Researchers may want to test this in their own settings with a truly scientific design for comparing parallel administrations, or rely on published literature to make the point. Given such support, it may be reasonable to rule out pure modality bias as a threat to validity.

Consider mixed-mode options

Cost, timing, and other advantages might lead to a decision to only use an online survey. Or, a number of mixed-mode options could be employed. For example, a pre-notification invitation with a web address might lead to 100 surveys being completed on the web in the first two days before phone calls begin, thereby saving the cost of 100 telephone completes.

There should be a serious consideration of design variables that will maximize response rates. The literature on mail surveys is rich with food for thought on this front. Some published studies have also experimented with the effects of varying some of these design variables in online surveys.

Don't substitute gimmicks for usable data and viable results

Finally, the actual technology for administration should be considered. Just as skips, rotations, and a host of other options exist for CATI surveys, web-survey technologies should be evaluated for their flexibility and ability to implement time-tested research practices. All of these capabilities should be applied with

principles of good research science. Bell-and-whistle technologies should never be substituted for the application of expertise in research sciences and research design.

Industry trends and history offer clues about the future of online surveys.

Questions about viability of web-based surveys should diminish as access to the web increases and issues of representativeness decrease. Acceptance should increase over time, as it did for telephone interviewing. At one time, telephone interviewing was viewed with skepticism since much research at that time was done door-to-door. No doubt web-based surveys are also a “disruptive” technology likely to change the face of how we collect data in the future.

Walker Information: at the forefront of next-generation survey methodology

There is a need for more truly experimental research on the pure modality effect that addresses the comparability issue without confounded designs. Walker Information is conducting this kind of testing, both collaboratively with the academic community and in applied organizational settings. If better designs continue to establish a lack of bias due exclusively to modality, then we can ask a new set of research questions. What survey design variables (incentives, pre-notification, sponsorship, personalization, length, visual appearance variables, scales, etc.) will help to maximize online participation and data quality?

Perhaps then we will also be turning our energy toward ways to capture “share of electronic attention” in a fast-moving world of dynamic interactive visual stimuli, constantly bombarding us and competing for our attention. Much like the advertising challenges that have come with hundreds of TV, satellite, and cable choices, perhaps competition will also skyrocket for web-based attention, including survey participation.

Conclusion: Web-based surveys are one very viable option

For now, novelty, speed, cost, individual convenience, and other advantages have brought web-based surveys rapidly into the research spotlight. It is unlikely that the next few years will bring any decrease in this trend. If anything, continued exponential growth can be expected.

With careful design consideration however, web-based surveys might well be the optimal applied methodology of choice in many situations. Having worked through all the trade-offs, they can be a practical, cost-effective, time-efficient, leading-edge approach to data collection that could change the face of research in today’s organizations.

Walker Information Internet Reporting: www.walkerinfo.com/wiir

Walker’s internet reporting tools combine sixty years experience in market research with innovative web technology. From data collection to results deployment, our comprehensive online offerings are fast, reliable, and secure. Call for a demonstration, or visit our site to learn more about Walker Information Internet Reporting: www.walkerinfo.com/wiir.



WiIR: Information at your fingertips



3939 Priority Way South Drive
Indianapolis, IN 46240-0972
Telephone: +1 317.843.3939
Toll free: +1 800.334.3939
Fax: +1 888.4WALKER
Email: info@walkerinfo.com
www.walkerinfo.com