

May 2010



The Fallacy of Online Surveys: No Data Are Better Than Bad Data

INTERNET OR ONLINE SURVEYS have become a popular and attractive way to measure opinions and attitudes of the general population and more specific groups within the general population. Although online surveys may seem to be more economical and easier to administer than traditional survey research methods, they pose several problems to obtaining scientifically valid and accurate results. A peer-reviewed article by Responsive Management staff published in the January-February 2010 issue of *Human Dimensions of Wildlife* details the specific issues surrounding the use of online surveys in human dimensions research. Reprints of the article can be ordered at <http://www.informaworld.com/smpp/content~content=a919147188~db=all~jumptype=rss>. Responsive Management would like to thank Jerry Vaske of Colorado State University for his assistance with the *Human Dimensions* article and for granting us permission to distribute this popularized version of the article.



Mark Damian Duda
Executive Director

Background

Natural resource and outdoor recreation professionals have found that gathering information through public opinion and attitude survey research gives them a precise and useful picture of what their organization's constituents think, need, and expect of them. Armed with this valuable information, they have been able to meet the future with organizational planning that is based on insight and knowledge obtained through scientifically valid, unbiased research methods.

It's a fact that conducting such research costs money. And in the current financial climate, with budgets being cut and uncertainty regarding what the future

holds, it makes sense for natural resource and outdoor recreation organizations to look for new ways to save money.

Online surveys are increasingly popular as an information-gathering tool. More and more online marketing companies offer online surveys at seemingly reasonable rates. Online surveys appear to be a great idea at first blush: they can be set up and administered in-house or contracted out, save time and money, and provide immediate results. But are online surveys a good idea? With few exceptions—the main one being employee surveys where every single employee has access to the Internet—for purposes of collecting

scientifically valid, accurate, and legally defensible data, the answer at this time is *no*. Recent research conducted by Responsive Management and published in the peer-reviewed journal *Human Dimensions of Wildlife* shows that online surveys can produce inaccurate, unreliable, and biased data. There are four main reasons for this: sample validity, non-response bias, stakeholder bias, and unverified respondents.

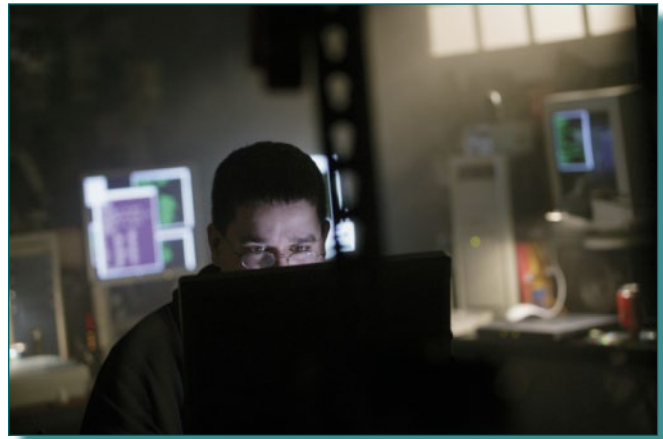
Sample Validity

For a study to be unbiased, *every member of the population under study must have an equal chance of participating*. When all members of the population under study have an equal likelihood of participating, probability sampling comes into play, and a relatively small sample size can yield results that accurately represent the entire population being studied.

For the most part, Internet surveys at this time cannot accomplish this, because there is no such thing as a representative sample of email addresses for various populations, including the general population and its subpopulations, such as registered voters, park visitors, or hunters and anglers. No “master list” of email addresses for any of these groups exists—not all people within these populations have an email address or access to the Internet. One exception is an online survey of a closed population in which *every member of that population has a verified email address and Internet access*. An internal survey of an organization in which all potential respondents are known and have guaranteed Internet access, usually through their workplace, is an example of this. Responsive Management has conducted this type of study (mainly employee surveys) for natural resource agencies in the past and has obtained results with scientifically valid sampling methodologies to back up study findings.

When online surveys are accessible to anyone who visits a website, the researcher has no control over sample selection. These *self-selected opinion polls* result in a sample of people *who decide to take the survey*—not a sample of scientifically selected respondents who represent the larger population. In this situation online survey results are biased because people who just happen to visit the website, people who are persuaded with a monetary or other incentive to sign up for the survey, people who have a vested interest in the survey results and want to influence them in a certain way, and people who are driven to the site by others are included in the sample. This results in a *double bias*, because this distortion is in addition to the basic sample already having excluded people who do not have Internet access.

Having access to a valid sample is the foundation for collecting data that truly represent the population being studied. Without a valid sample, every bit of data obtained thereafter is called into question.



Non-Response Bias

Non-response bias in online surveys is complicated by the most egregious form of self-selection. People who respond to a request to complete an online survey are likely to be more interested in or enthusiastic about the topic and therefore more willing to complete the survey, which biases the results. In fact, the very nature of the Internet, as an *information-seeking* tool, contributes to this form of bias. For example, if someone who is interested in the subject matter of a survey uses a search engine, such as Google, to seek out information on the subject, that person is more likely to find an online survey on that topic. In this way, more people with a heightened interest in the topic are driven to the online survey.

With a telephone survey, people are contacted who are not necessarily interested in the topic, and if they are not enthusiastic about completing the survey, a trained interviewer can encourage them to do so despite their disinterest, leading to results that represent the whole population being studied, not just those with an interest in the subject.

Another contributor to non-response bias in online surveys is spam and unsolicited mail filters. Users can set the degree of message filtering, and if the tolerance is set strictly enough, they may not even see a request to participate in an online survey because the filter will automatically “trash” the email request when it is delivered. This removes these individuals from the possibility of receiving an invitation to participate in an online survey.

Potential respondents to an email request to participate in an online survey may have more than one, and sometimes multiple, email addresses. It is impossible to know which is the primary address for an individual or even if the person checks the account on a regular basis for incoming mail.

Stakeholder Bias

Unless specific technical steps are taken with the survey to prevent it, people who have a vested interest

in survey results can complete an online survey multiple times and urge others to complete the survey in order to influence the results. This is a common occurrence, especially regarding issues that elicit high levels of concern, such as, in the fish and wildlife context, when an agency wants to measure opinions on proposed regulation changes. Some Internet-savvy individuals have even written automated programs that repeatedly cast votes to influence a poll's results.

Even when safeguards against multiple responses are implemented, there are ways to work around them. If there is a protocol in place that limits survey completions to one per email address, it's easy to go online and open a new email account with a new address and then complete another survey through that email address. If access is limited to one survey completion per computer, completing another survey can be done on a separate computer, at a friend's home, in the workplace, or in a public library, for example. And in the case of online surveys where individuals have to sign up in order to participate, they can sign up under multiple names and email addresses and participate multiple times through each of those email addresses.

Unverified Respondents

Because of the inability to control who has access to online surveys, there is no way to verify who responds to them—who they are, their demographic background, their location, and so on. As stated earlier, even when safeguards are implemented to control access to online surveys, there are multiple ways to circumvent those safeguards.

A complicating issue is when an organization offers incentives for completing online surveys. Whether it's a chance to win a prize, discounts on purchases, a gift certificate, or some other benefit, offering an incentive without having close control over the sample simply encourages multiple responses from a single person. If someone has a strong desire to win the item, he or she can find ways around any safeguards against multiple responses and complete several surveys, thereby increasing his or her chances of winning the item.

Examples

Three recent collaborative projects with state fish and wildlife agencies gave Responsive Management an opportunity to compare the results of online versus scientific telephone surveys within the same study topics.

North Carolina Sunday Hunting Study

Sunday hunting has been a controversial issue in North Carolina, with strong feelings among both supporters and opponents. To better understand the issue, the North Carolina Wildlife Resources Commission (NCWRC), Virginia Tech, and Responsive Management



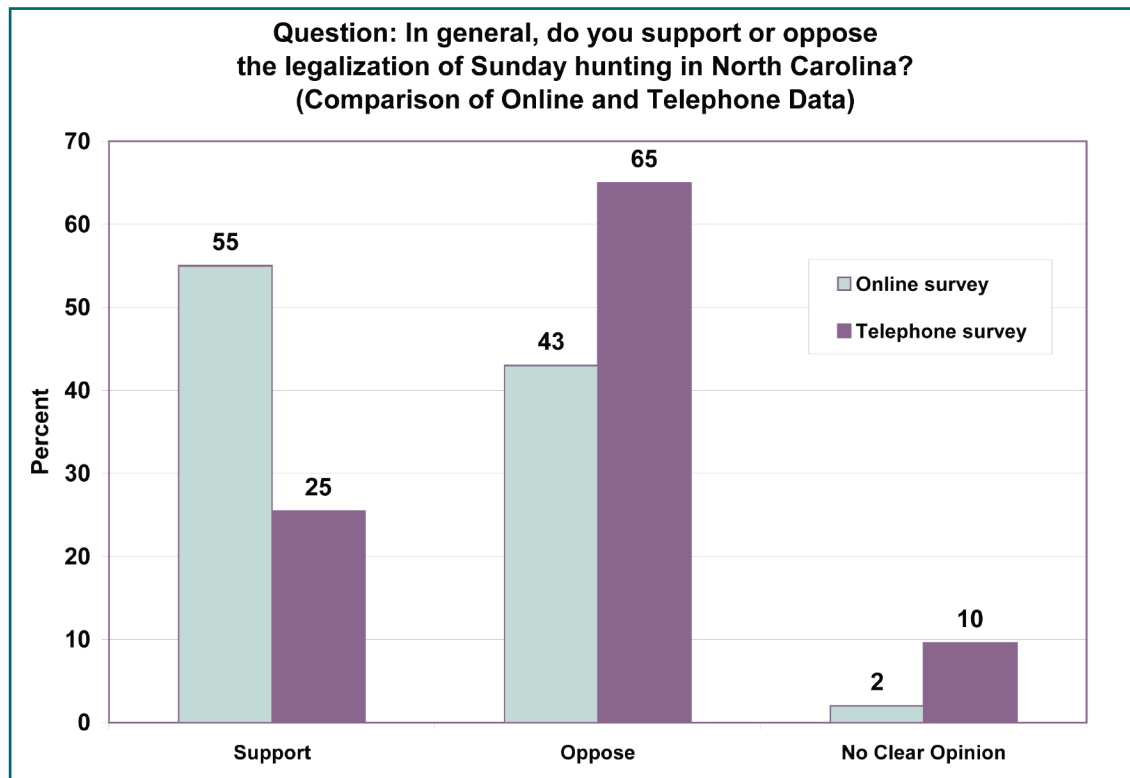
collaborated on a study to assess public opinion on Sunday hunting. The study consisted of an online opinion poll, a telephone survey, and an economic analysis.

The online poll was developed and placed on the NCWRC website to elicit feedback on support or opposition to Sunday hunting. The online poll was developed primarily as an outlet for people who wanted to be heard. At the same time, a scientific telephone survey was conducted by Responsive Management, Virginia Tech, and the NCWRC.

The results of the two surveys were markedly different. The online poll showed that 55% of respondents supported Sunday hunting, whereas 43% opposed it, and 2% had no clear opinion. The telephone survey showed that 25% of respondents supported Sunday hunting, whereas 65% opposed, and 10% had no clear opinion. These differences are well outside of any acceptable margin of error for a valid study.

The telephone survey, because it used a randomly generated sample of North Carolina residents, accurately reflected the opinions of the population as a whole. Because more than 1,000 individuals were interviewed, the sampling error was at most plus or minus 2.815 percentage points.

Far more people in the telephone survey of North Carolina residents opposed Sunday hunting compared to those who responded to the online poll. Only 25% of the telephone respondents supported Sunday hunting, whereas 55% of those who responded to the online poll supported it. The telephone survey found a fivefold increase as compared to the online poll in people who had no clear opinion on the subject of Sunday hunting. This indicates that far more people with a vested interest in the results completed the online poll; when the general population was scientifically surveyed, a truer number of North Carolinians who had no clear opinion was revealed. In short, had the NCWRC gone with the online poll results, it would have gotten an inaccurate read on what the public was thinking regarding Sunday hunting in the state.



“While I was not surprised that there were differences between the online interface and telephone survey results, given that the telephone survey used probability sampling and anyone who chose to could give their opinion online, I was somewhat surprised at the size of these differences,” said Dain Palmer, Human Dimensions Biologist at the NCWRC.

Arizona Big Game Hunt Permit Tag Draw Study

In 2006 the Arizona Game and Fish Department (AZGFD) conducted an online survey to assess hunter attitudes toward the Arizona Big Game Hunt Permit Tag Draw, a topic with a high degree of interest to Arizona hunters. When the data collection for the online survey was completed, the AZGFD had doubts about its accuracy and worked with Responsive Management to conduct a non-response bias analysis. A telephone survey of the online survey non-respondents was conducted to assess non-response bias. In other words, those who were contacted by email but who did not respond were contacted by telephone and interviewed.

For the online survey, a link to the survey site was distributed by email to individuals who had provided an email address when applying for the 2006 Fall Big Game Draw. Duplicate and invalid email addresses were removed, and the survey was sent to a total of almost 60,000 Fall Big Game Draw applicants.

The online survey included a unique website address for each email address, which “closed” the survey to that respondent once he or she completed it. This

ensured that multiple responses from a single email address did not occur and that a response from a specific email address could be tracked if necessary. For the telephone survey, people who did not respond to the email request were contacted and interviewed.

Responsive Management analyzed those who responded to the survey and those who did not and identified several statistically significant differences between the groups. Of the 766 variables analyzed in the study, differences for 312 variables—41% of the variables analyzed—were statistically significant. This means that, on almost half of the variables where those who responded to the online survey were compared to those who did not respond, there was a meaningful difference between how they responded to the same question.

Why are these differences a problem? *Simply because they exist.* If both of these surveys were representative of the population group under study—Arizona hunters who applied for the 2006 Fall Big Game Draw and provided an email address—there would be *no statistically significant differences* between how the people who responded to the email request answered the questions and how those who did not respond to the request answered the questions. (This bias is in addition to the basic bias of omitting people who did not provide an email address when applying, as described in more detail in the South Carolina study discussed below.)

“Our initial reaction to the big game hunt permit study was that it validated what we had been hearing

anecdotally for a long time from the general hunting community,” said Ty Gray, an Assistant Director with the AZGFD. “Specifically, getting to go hunting or getting a permit tag were very important factors which both groups (Web and phone respondents) agreed on. However, as we started to look closer at some of the other variables, we saw that there were differences that indicated some bias with the online survey—among those was who was more likely to respond to it.”

Again, if this were a valid sample to begin with, there would be no statistically significant differences between these two groups. In short, there were major differences in responses, with the online survey providing biased and inaccurate data.

“Game and Fish commissioners regularly have to make important decisions under extreme pressure from special interest groups,” said Bob Hernbrode, former chairman of the Arizona State Fish and Game Commission. “Valid social science such as this Arizona study often suggests significantly different outcomes than special interest input would suggest. We need to understand the potential of poorly designed studies and such things as non-response bias.”

South Carolina Saltwater Fishing and Shellfishing Study

In 2009, Responsive Management and the South Carolina Department of Natural Resources (SCDNR) collaborated on a survey to assess participation in and opinions on saltwater fishing and shellfishing in South Carolina and to better understand the accuracy and potential of online surveys. Two different methodologies were used: a scientific survey conducted by telephone and a survey conducted via the Internet. This study is a best-case scenario regarding online surveys because it involved a closed population—people who obtained a South Carolina Saltwater Recreational Fisheries License. If online surveys could produce accurate data, this would be the study that would prove it.

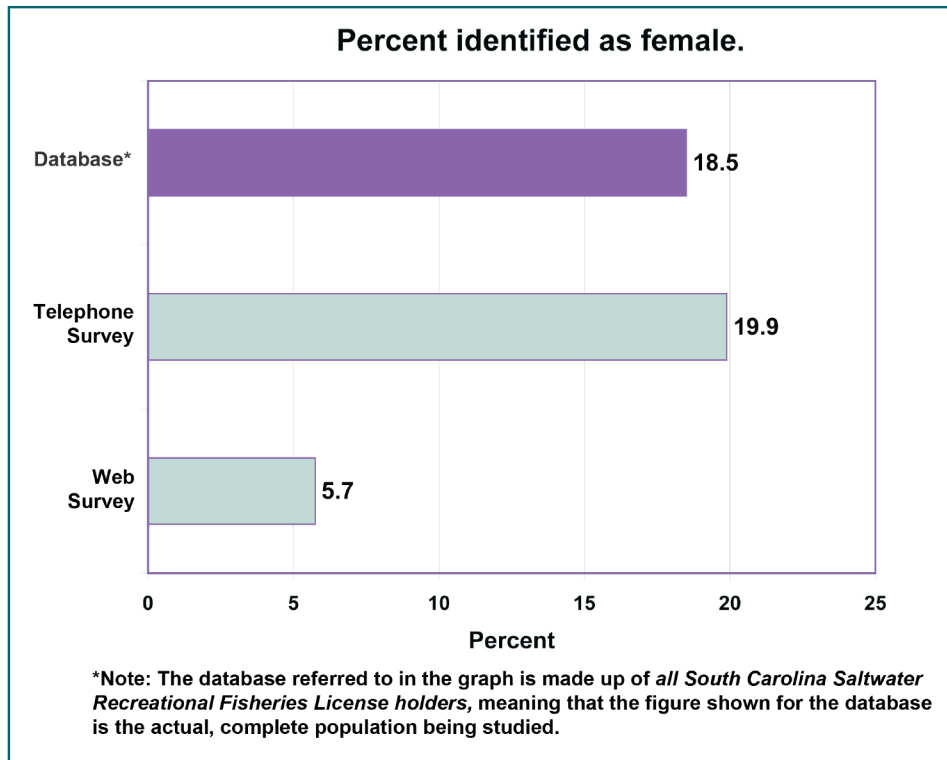
The researchers were able to test this because they had a base sample—the entire database of Saltwater Recreational Fisheries License holders, including demographic and geographic information for each license holder—that could be compared to both the telephone and online survey results. When the two methodologies were compared, the telephone survey yielded results that accurately reflected the entire population, whereas the online survey did not. This is because the telephone survey included a greater proportion of the population under study than the online survey did. The telephone survey sample was randomly drawn from the *entire population* of people who held a Saltwater Recreational Fisheries License; for license holders who did not provide a telephone number, their telephone number was identified by

reverse lookup. Therefore, every license holder had an equal chance of being contacted by telephone to take part in the survey. The online survey used a sample consisting of people who held Saltwater Recreational Fisheries Licenses *who provided an email address when they purchased their licenses*. This systematically excluded license holders who did not have computer access and license holders who chose not to provide an email address. While one might think this is not important, the results showed otherwise. Because of the systematic exclusion of these license holders, the results of the online survey were inaccurate from the outset.

The information from the database indicated that, out of a total population of 103,000 license holders, the online survey had an original sample of approximately 16,100 license holders with email addresses, which produced 12,405 license holders in the sample after email addresses that were undeliverable were removed. Therefore, even before any contacts were made, the online survey had eliminated approximately 88% of the possible sample, and did so in a systematic way, which is the very definition of bias. In addition, there was a notable non-response bias: of the 12,405 license holders contacted by email, only 2,548, or 20.5%, responded to the online survey. These problems lead to a double bias: first, the exclusion of people with no email address, and second, exclusion of those who did not respond to the online survey.

If both of these surveys were representative of the population group under study . . . there would be no statistically significant differences between how the people who responded to the email request answered the questions and how those who did not respond to the request answered the questions.

With a scientifically selected sample, reducing the sample size to this degree would not be a problem, because the smaller sample would be representative of the population as a whole—the methodology used to select the sample from the total population being studied would ensure that this would be the case, within a demonstrable sampling error. But in the case of a sample that is not scientifically generated, reducing the sample size in this way simply would bias the results even more—the more the sample is reduced, the more biased it becomes.



Because they had access to the database of all license holders, which included demographic and geographic information, Responsive Management statisticians were able to determine that, from the outset, the respondents who provided email addresses were different from the sample as a whole. If the online sample had been valid, there would have been no statistically significant differences between the two—each sample would have been consistent with

This study is a best-case scenario regarding online surveys because it involved a closed population—people who obtained a South Carolina Saltwater Recreational Fisheries License. If online surveys could produce accurate data, this would be the study that would prove it.

and representative of the population as a whole: the 103,000 license holders being studied.

When the online survey was completed and the data were analyzed, the online survey respondents were found to be, in general, a more educated and affluent group, and were also disproportionately male. Of particular note, 5.7% of the online survey sample was female, whereas 19.9% of the telephone sample

was female; in reality, 18.5% of the license holder database—the actual number of license holders—was female. The telephone results were therefore much closer to the truth than the online results. In fact, the online results were so far off the mark that they would have led to highly inaccurate findings, because females were not represented in the proportion that they should have been.

“When we initially saw the differences between the online and telephone surveys, we were not too surprised that the results differed, simply due to the fact that only a small portion of license holders provide their email address on their saltwater fishing license application,” said Julia Byrd of the SCDNR’s Office of Fisheries Management. “Due to this, we thought the results of the online survey might be biased because certain demographic groups would be over- or underrepresented. This was shown in the results.”

The Result

As a result of these problems, obtaining representative, unbiased, scientifically valid results from online surveys is not possible at this time, except in the case of the closed population surveys, such as with employee surveys, described earlier. This is because, from the outset, *there is no such thing as a complete and valid sample*—some people are systematically excluded, which is the very definition of bias. In addition, there is no control over who completes the survey or how many times they complete the survey. These biases increase in a stepwise manner, starting out with the basic issue of

excluding those without Internet access, then non-response bias, then stakeholder bias, then unverified respondents. As each of these becomes an issue, the data become farther and farther removed from being representative of the population as a whole.

For a more detailed look at these examples and more information on the drawbacks of online surveys

in the context of human dimensions research, see Duda, M.D, & Nobile, J.L., "The Fallacy of Online Surveys: No Data Are Better Than Bad Data," *Human Dimensions of Wildlife* 15(1): 55–64. Reprints of the article can be ordered at <http://www.informaworld.com/smpp/content~content=a919147188~db=all~jumptype=rss>.

What Constitutes a Valid Survey?

Just a few years ago, when the results of online surveys were displayed, a disclaimer would accompany the findings: "Not a scientific survey." Unfortunately, while the inherent biases of online surveys remain, this cautionary note has been dropped for the most part, or at least relegated to a minor position below the survey results, in small type.

This is unfortunate, as most online surveys are not scientific, produce inaccurate data, and do not reflect the opinions or attitudes of the population under study. In short, most online surveys produce erroneous data that can lead to erroneous conclusions and subsequently to bad decision-making by agencies, organizations, and businesses. It's better to have no data than inaccurate data.

There are two general types of sampling that one can use to study populations: (1) probability sampling, and (2) non-probability sampling. *Probability sampling* is sampling that is scientifically generated and where every member of the population being studied has an equal chance of participating, and where that probability can be demonstrated mathematically. These are the type of samples used by polling organizations such as Gallup, as well as by research organizations that want a representative sample of a specific population for quantitative study, such as Responsive Management. *Non-probability sampling* means that participants are selected without benefit of a scientifically valid sampling plan. These include surveys where people volunteer to participate or surveys where participants are chosen without a statistically valid sample population.

Online surveys are largely conducted through non-probability sampling: access to the survey is not controlled, and anyone can participate. The Internet usually features three kinds of non-probability surveys. The first consists of online polls or surveys in which anyone can participate. These are sometimes referred to as *self-selected opinion polls*, or SLOP surveys, meaning that people *who decide to take the survey* make up the sample. The second type is closed population surveys, where a common factor exists among the respondents, but respondents are *still* self-selected within that population, and access to the survey is not necessarily controlled. The third is specific closed population surveys, in which there is a specific group of people, more control over access, and some email representation.

The problem with all of these methods is present before even one survey question is asked: *lack of a valid sample*. Even if a statistically defensible population is generated, once the survey is placed in the context of the Internet, it is not representative of the population as a whole. This is because people without Internet access—still a large segment of the U.S. population—are systematically excluded from the sample, as are people who have online access but do not see the survey online. Notwithstanding other problems, such as lack of control over who answers the survey and how many times they do so, this basic problem of sample invalidity remains.

One way to determine whether survey results are valid is to see whether the organization that conducts the survey discloses its sampling methodology. If the organization does not reveal this information or offers incomplete information, chances are that the results are not valid. A scientific study must adhere to strict methodology in order for results to be accurate.

No amount of weighting can make up for an unrepresentative sample. Weighting in this context is an attempt to create a more accurate picture than the data will allow. The result is "photoshopped" survey results—the information looks nice and seems to be complete, but it isn't representative or scientifically valid. No matter what you do with the data, the fact remains that you cannot make it any better, any "more," than what it was in the first place.