Human Dimensions of Wildlife, 15:55–64, 2010

ISSN: 1087-1209 print / 1533-158X online
DOI: 10.1080/10871200903244250

Routledge
Taylor & Francis Group

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

MARK DAMIAN DUDA AND JOANNE L. NOBILE

Responsive Management, Harrisonburg, Virginia, USA

Internet or online surveys have become attractive to fish and wildlife agencies as an economical way to measure constituents' opinions and attitudes on a variety of issues. Online surveys, however, can have several drawbacks that affect the scientific validity of the data. We describe four basic problems that online surveys currently present to researchers and then discuss three research projects conducted in collaboration with state fish and wildlife agencies that illustrate these drawbacks. Each research project involved an online survey and/or a corresponding random telephone survey or non-response bias analysis. Systematic elimination of portions of the sample population in the online survey is demonstrated in each research project (i.e., the definition of bias). One research project involved a closed population, which enabled a direct comparison of telephone and online results with the total population.

Keywords Internet surveys, sample validity, SLOP surveys, public opinion, non-response bias

Introduction

Fish and wildlife and outdoor recreation professionals use public opinion and attitude surveys to facilitate understanding their constituents. When the surveys are scientifically valid and unbiased, this information is useful for organizational planning. Survey research, however, costs money. Given the current budgetary climate and the uncertainty of the future, organizations are looking for ways to save money. Strategic planning and human dimensions information-gathering are no exception.

Online surveys are becoming increasingly popular as information-gathering tools. Marketing companies offer online surveys at seemingly reasonable rates. Online surveys *appear* to be easy to set up and administer in-house, can save time and money, and provide immediate results. Unfortunately, online surveys seldom provide scientifically valid, accurate, and legally defensible data. Recent collaborative research conducted by Responsive Management and three state fish and wildlife agencies provides clear examples of how online surveys can produce inaccurate, unreliable, and biased data. There are four main reasons for this: (a) sample validity, (b) non-response bias, (c) stakeholder bias, and (d) unverified respondents.

The challenges that online surveys present to obtaining scientifically valid survey results have been pointed out by others. Dillman, Smith, and Christian (2009), for example, cite the lack of standardization regarding e-mail address structure and how e-mail addresses are created, the absence of an online equivalent to the random digit dialing (RDD) algorithm for random selection of telephone numbers, and respondents' varying

Address correspondence to Mark Damian Duda, Responsive Management, 130 Franklin Street, Harrisonburg, VA 22801, USA. E-mail: mark@responsivemanagement.com

levels of capability in using the Internet as reasons that online surveys can result in invalid probability samples. Because a statistically valid sample frame cannot be established for most online surveys, non-response bias cannot be determined. As a result, little is currently known about non-response bias in Web-based surveys (Couper, 2000). Lack of control over the survey sample caused by stakeholder bias and unverified respondents can compromise the validity of online surveys and has been discussed in the literature (Bethlehem, 2008; Couper, 2000; Dillman et al., 2009; Vaske, 2008; Zhang, 1999).

Although there are several types of online surveys, including e-mails with embedded surveys, e-mail message requests for participation that include a link to a Web-based survey, and "open" Web-based surveys with no related e-mail request for participation, the latter two types of surveys are discussed here.

Sample Validity

Online surveys can yield inaccurate information regarding *sample validity.* For a study to be unbiased, *every member of the population under study must have a known chance of participating.* Unfortunately, it is not possible to draw a general public sample, where every member of the population has a known chance of being selected, from a population of online users. There are no representative samples of e-mail addresses for various populations, including the general population and, for example, registered voters, park visitors, and hunters and anglers. A "master list" of e-mail addresses for these groups does not exist because not all people within these populations have e-mail addresses or access to the Internet. When all members of the population under study have a known likelihood of participating, probability sampling allows for a relatively small sample size that can accurately represent the entire population (Dillman et al., 2009; Vaske, 2008).

Exceptions, however, can occur. One example is an internal survey of an agency or organization in which *all* potential respondents are known and have guaranteed Internet access, usually through their workplace. Studies such as these of fish and wildlife agencies have obtained results with scientifically valid sampling methodologies (e.g., Responsive Management, 2006b, 2007). Another exception occurs when participants are contacted by telephone and given the option of responding to the survey on the Internet, as opposed to a telephone interview. In these *mixed-mode surveys* (Couper, 2000; Dillman et al., 2009), the sample is not affected; respondents can still be selected based on randomly generated telephone numbers or a valid, representative sample of the population under study.

Online surveys can also result in invalid samples through *self-selection.* When online surveys available through an organization's Web site are accessible to anyone who happens to visit the site, the researcher has no control over sample selection. These are sometimes referred to as *self-selected opinion polls,* or SLOP surveys (American Association of Public Opinion Researchers, 2009). Such online survey results are biased because the respondents were not scientifically selected. The participants: (a) just happened to visit the Web site, (b) were persuaded by monetary or other incentives, (c) have a vested interest in the results or want to influence the findings, and/or (d) were driven to the site by others in the sample. This results in a *double bias* because the sample already excluded people without Internet access. People who are unaware that the survey exists do not have any chance of participating, which introduces further sampling bias. Access to a valid sample is the foundation for collecting data that are representative of the population. Without a valid sample, all data are questionable (Vaske, 2008).

Non-Response Bias

Research has shown that people who do not respond to requests to participate in a survey have the potential to be different from those who do respond, especially in general population surveys where response rates are low (Vaske, 2008). In these cases, *non-response bias* must be accounted for to ensure that survey results are representative of the population being studied. Because of the basic sample validity problems that online surveys present, a statistically valid sample frame usually cannot be established, and therefore the presence or absence of non-response bias cannot be determined.

There are several ways unique to online surveys in which respondents and non-respondents may differ. People who respond to a request to complete an online survey are likely to be more interested in the topic and therefore more willing to complete the survey, which biases the results. The very nature of the Internet, as an information-seeking tool, contributes to this form of bias. For example, someone who is interested in a survey's subject matter is more likely to seek information on the subject (e.g., via Google), and is more likely to stumble upon the online survey. 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 more closely represent the whole population being studied, not just those with an interest in the subject. When a potential respondent is asked in an e-mail message to complete a survey, it is easy to delete the message (Vaske, 2008). Thus, those with a special interest in the topic will be more likely to respond, whereas those with only mild or no interest at all will be less likely to respond.

Mail filtering contributes to non-response bias in online surveys when e-mails are blocked or discarded by the recipient. Recipients can set the degree of message filtering. When this tolerance is strict, requests to participate may not be seen because the filter automatically "trashes" the survey request when it is delivered. Some users set up their e-mail account such that messages are not received unless the sender has been pre-approved. This completely removes these individuals from the possibility of receiving an invitation to participate in an online survey.

Finally, some potential respondents may have multiple e-mail addresses. It is impossible to know which is the individual's primary address or even if the person checks the account on a regular basis for incoming mail.

To illustrate the impact of non-response bias, two studies are presented. The first study is a non-response bias analysis conducted by Responsive Management and the Arizona Game and Fish Department (Responsive Management, 2006a). The second is a saltwater fishing and shellfishing study conducted by Responsive Management and the South Carolina Department of Natural Resources (Responsive Management, 2009a, 2009b).

Stakeholder Bias

Stakeholder bias can affect online survey results. People with a vested interest in the survey results can complete an online survey multiple times and urge others to complete the survey in order to influence the results. This practice (a.k.a. *poll crashing*) can be especially prevalent regarding issues that elicit high levels of concern (e.g., when a fish and wildlife agency wants to measure opinions on proposed regulation changes). Internet-savvy

individuals can write automated programs (i.e., *Internet bots*) that repeatedly cast votes to influence a poll's results.

Even when safeguards against multiple responses are implemented, there are workarounds. If a software protocol limits survey completions to one per e-mail address, a new e-mail account with a new address can be created and another survey completed with that address. If access is limited to one survey completion per computer, other surveys can be completed on other computers (e.g., at a friend's home, the workplace, or a public library). With online surveys where individuals have to sign up in order to participate, respondents can sign up under multiple names and e-mail addresses and vote multiple times through each address.

An example of stakeholder bias occurred in 2008, shortly after the Republican presidential candidate, Arizona Senator John McCain, announced that his vice presidential running mate would be Alaska Governor Sarah Palin. The Public Broadcasting Service (PBS) posted a one-question online poll on its Web site that asked, "Do you think that Sarah Palin is qualified to serve as Vice President of the United States?" with a choice of "Yes" or "No" as an answer (Siceloff, 2008). Respondents were initially allowed to vote repeatedly, but due to an overwhelming response, PBS implemented a registration system that limited votes to one per computer. Even after attempts to control multiple voting and moving the poll to a less prominent PBS Web page, people still sought out the site to cast votes. By October 15, 2008, 52 million votes were cast; 49% answered "Yes" and 49% answered "No" (Siceloff, 2008). In situations involving an issue that people feel strongly about, online survey results are not reliable because of the tendency to respond repeatedly.

Unverified Respondents

Without controls on who accesses online surveys, there is no way to verify respondents' demographic backgrounds. Even with safeguards implemented to control access to online surveys, there are multiple ways to circumvent those safeguards. The issue is further complicated when incentives are offered for completing online surveys (e.g., a chance to win a prize, discounts on purchases, a gift certificate). Incentives without close control over the sample simply encourage multiple responses from a single person. People with a strong desire for a prize or item can find ways around multiple response safeguards and complete several surveys, thereby increasing their chances of winning.

Weighting Data

For telephone and mail surveys, if the results of a non-response bias check show differences between people in the sample who completed a survey and those who did not complete it, data may need to be weighted for the sample results to be representative of the target population (Vaske, 2008). Weighting may also be necessary if the population proportions are known in advance and survey results reveal that specific groups or segments are overrepresented or underrepresented. Such weighting strategies, however, assume probability sampling. No amount of weighting can make up for an unrepresentative sample such as is typically found in online surveys.

Empirical Illustration

Three recent projects by Responsive Management and state agencies are used to compare results of online versus scientific telephone surveys within the same study topics.

For each telephone survey, a central polling site at Responsive Management allowed for quality control over the interviews and data collection. The survey center managers and other professional staff conducted project briefings with interviewers prior to administration of the telephone surveys. Interviewers were instructed on type of study, study goals and objectives, handling of survey questions, interview length, termination points and qualifiers for participation, interviewer instructions within the survey instrument, reading of the survey instrument, skip patterns, and probing and clarifying techniques necessary for specific questions on the survey instruments. Pre-tests of the questionnaires were conducted and revisions were made to the questionnaires based on the pre-tests.

North Carolina Sunday Hunting Study

Sunday hunting has been controversial in North Carolina, with strong feelings among both supporters and opponents. To help the North Carolina Wildlife Resources Commission (NCWRC) better understand the issue, the agency worked with Responsive Management and Virginia Tech to assess public opinion on Sunday hunting (Responsive Management, 2006c; Virginia Polytechnic and State University/Responsive Management, 2006). The study included an online opinion poll and a telephone survey.

The online poll was developed and placed on the NCWRC Web site to elicit feedback on support or opposition to Sunday hunting. The online poll (a non-random sample) was developed primarily as an outlet for people who wanted to be heard. In addition to a section asking about the respondent's gender, residence, land ownership status, and hunting participation, the online poll included an open-ended comments section. Responses were accepted from January 31, 2006, to August 31, 2006.

More than 10,000 responses to the poll were received through the NCWRC Web site. Responses were also collected by e-mail, telephone, and mail. All responses by all communication modes were categorized by (a) support for legalization of Sunday hunting in general, (b) opposition to legalization of Sunday hunting in general, and (c) no clear opinion expressed on the subject.

Responses were reviewed to eliminate multiple respondents; a small percentage of submissions was discarded based on strong evidence that the individuals responded to the poll multiple times. A few submissions were discarded because they were unrelated to the topic of Sunday hunting. The online poll, however, was not representative of the views of the general population or hunters because only those who were aware of the issue submitted their comments online or in writing (Virginia Polytechnic and State University/ Responsive Management, 2006).

For the telephone survey, interviews were conducted Monday through Friday from 9:00 a.m. to 9:00 p.m., Saturday from noon to 5:00 p.m., and Sunday from 5:00 p.m. to 9:00 p.m., local time, using a five-callback design to maintain the representativeness of the sample, to avoid bias toward people easy to reach by telephone, and to provide an equal opportunity for all to participate. When a respondent could not be reached on the first call, interviewers placed subsequent calls on different days of the week and at different times of the day. The survey was conducted in June and July 2006 (n =1,212 completed interviews).

Findings of the telephone survey were reported at a 95% confidence interval. For the entire sample of North Carolina residents, the sampling error was at most $\pm$2.815 percentage points. Sampling error was calculated using the formula detailed in Dillman (2000, p. 206).

The overall results indicated marked differences between the responses generated by the online poll and the scientific telephone survey results. Over half (55%) of the online

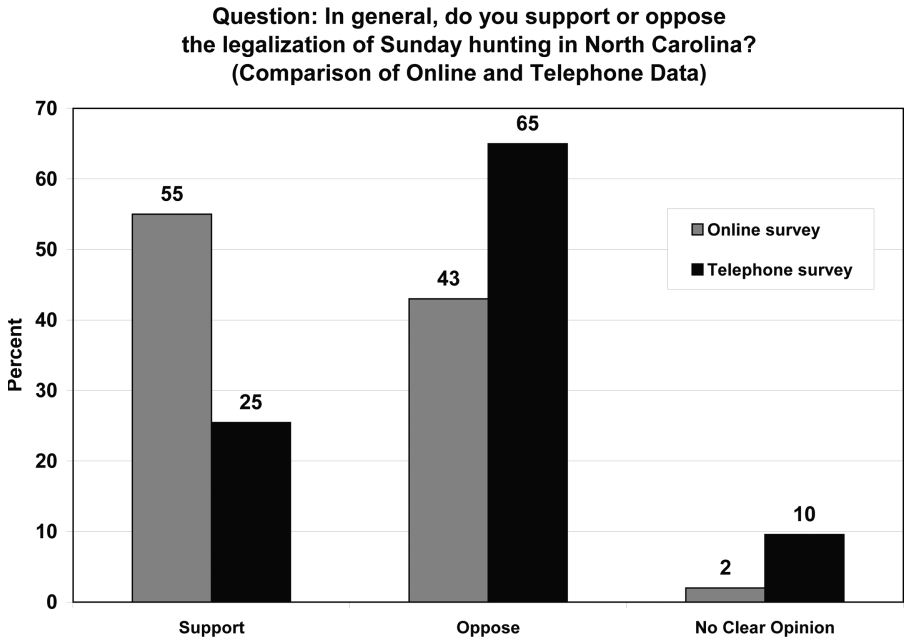

Figure 1. A comparison of results from a 2006 telephone survey conducted by Responsive Management, Virginia Tech, and the North Carolina Wildlife Resources Commission with an online survey posted on the NCWRC Web site (Responsive Management, 2006c).

respondents supported Sunday hunting, 43% were opposed, and 2% had no clear opinion (Figure 1). A quarter (25%) of the telephone respondents supported Sunday hunting, 65% were opposed, and 10% had no clear opinion. The difference (10% telephone vs. 2% online) in people who had no clear opinion on the subject of Sunday hunting suggests that more people with a vested interest in the results completed the online poll. Relying on the online poll results would have led to false impression of what the public was thinking regarding Sunday hunting in the state. Because the telephone survey used a randomly generated sample of North Carolina residents, results accurately reflected the population.

Arizona Big Game Hunt Permit Tag Draw Study

In 2006 the Arizona Game and Fish Department (AGFD) 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. A 53-question online survey was developed cooperatively by the AGFD and Responsive Management. The survey was administered online by zoomerang.com, an online survey vendor. A hyperlink to the Web survey was distributed by e-mail to individuals who had provided an e-mail address when applying for the 2006 Fall Big Game Draw. Duplicate and invalid e-mail addresses were removed, and the survey was sent to 59,967 Fall Big Game Draw applicants. The Web survey program created a unique Web address for each e-mail address. The unique Web address ensured that multiple responses from a single e-mail address were removed, and a response from a specific e-mail address could be tracked if necessary. The Web survey was conducted August 15 to September 10, 2006. A total of 15,156 completed Web questionnaires was obtained.

Despite these methodological safeguards, the AGFD had doubts about the online survey's 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 (Responsive Management, 2006a). Those who were contacted by e-mail but who did not respond were contacted by telephone and interviewed.

The non-response telephone survey questionnaire emulated the online questionnaire and was developed cooperatively by the AGFD and Responsive Management. Interviews were conducted using the procedures outlined earlier. The non-response telephone survey was conducted in October 2006. A total of 202 completed interviews was obtained. Findings of the non-response telephone survey are reported at a 95% confidence interval. For the entire sample of non-respondents, the sampling error was at most ±6.88 percentage points.

Those who responded to the original e-mail request and completed the online survey differed statistically from the non-responders who were contacted in the non-response telephone survey on 312 of the 766 variables examined in the study (i.e., 41% of the variables). Such differences are a problem simply because they exist. If both surveys were representative of the population of Arizona hunters who applied for the 2006 Fall Big Game Draw and provided an e-mail address, there should have been *no statistically significant differences.* This bias is in addition to the basic bias of omitting people who did not provide an e-mail address when applying.

Non-responders to the online survey were older on average than those who did respond (χ^2 = 668.50, df = 68, $p < .001$). Non-responders to the online survey were also more likely than those who did respond to rate the importance of receiving two big game hunt permit tags each year as a 10 (extremely important) (χ^2 =10.32, df =1, $p < .001$), to be willing to continue applying for a big game permit tag without success for longer periods of time (e.g., more than 10 consecutive years) (χ^2=143.69, df = 1, $p < .001$), and to not belong to or have donated to hunting or conservation organizations (χ^2 = 14.06, df = 1, $p < .001$).

South Carolina Saltwater Fishing and Shellfishing Study

In 2009, Responsive Management and the South Carolina Department of Natural Resources (SCDNR) conducted 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 (Responsive Management, 2009a, 2009b). 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 the evaluation of online surveys because it involved a closed population (i.e., people who obtained a South Carolina Saltwater Recreational Fisheries License). The entire database of Saltwater Recreational Fisheries License holders, including demographic and geographic information for each license holder, could be compared to both the telephone and online survey results.

The telephone survey was developed cooperatively by Responsive Management and the SCDNR. Interviews were conducted as described previously. The survey was conducted January 19 through January 23, 2009. A total of 1,709 completed interviews was obtained. Findings of the telephone survey are reported at a 95% confidence interval. For the entire sample, the sampling error was ±2.35 percentage points.

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 numbers were identified by reverse lookup. 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 e-mail 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 e-mail address.

Out of the total population of 102,610 license holders, approximately 16,100 license holders had provided e-mail addresses, which produced a sample size for the online survey of 12,405 license holders after undeliverable e-mail addresses were removed. Thus, even before any contacts were made, the online survey had eliminated approximately 88% of the possible sample systematically (i.e., the definition of bias). There was also notable non-response. Of the 12,405 license holders contacted by e-mail, only 2,548 responded online.

With a scientifically selected probability sample, reducing the sample size to this degree would not be a problem. The smaller sample would still be representative of the population as a whole, within a demonstrable sampling error. For the non-random online sample, reducing the sample size in this way biased the results.

With access to the database of all license holders, we were able to determine from the outset that respondents who were reachable by e-mail (n=12,405) differed statistically from the total population of license holders (n = 102,610) on several key demographic and geographic variables. Had the online sample been representative, there would have been *no statistically significant differences*.

The online respondents were more educated (χ^2= 42.23, df =1, $p < .001$), affluent (χ^2 = 75.15, df = 1, $p < .001$), and disproportionately male (χ^2 = 141.77, df = 2, $p < .001$) compared to the total population of license holders. About 6% of online respondents were female, compared to about 19% of the total population of license holders (Figure 2).

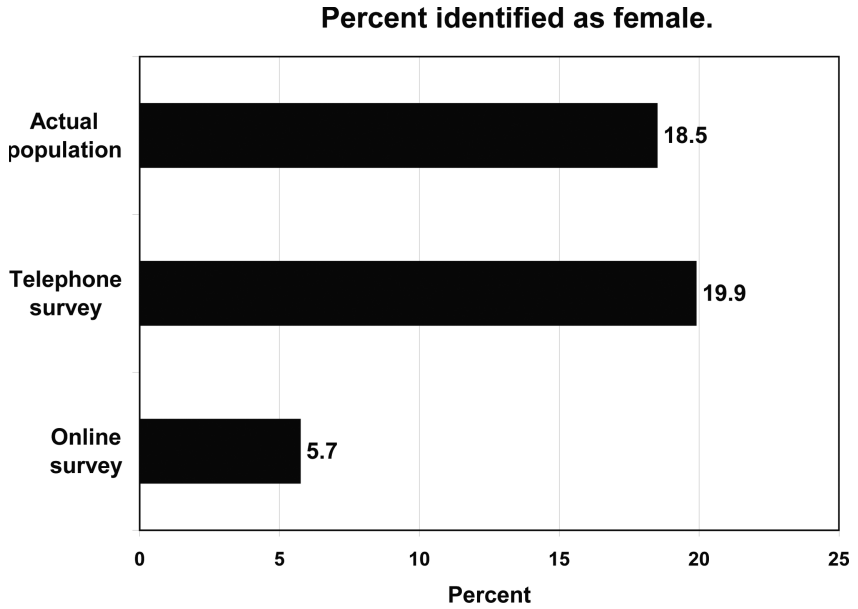

Figure 2. In an online survey regarding saltwater fishing and shellfishing in South Carolina, women were represented at only one third of what they should have been to represent the population as a whole. The telephone survey was much closer to representing the true proportion of females in the actual population (Responsive Management, 2009a).

By comparison, the telephone survey results were consistent with all the variables examined. For example, 19.9% of the telephone respondents were female, compared to 18.5% of the total population of license holders (Figure 2), well within an acceptable margin of error. Thus, the telephone results reflected the total population more accurately than the online results.

The online survey had a higher percentage of respondents in the "Don't know/no response" category than the telephone survey, and many questions were not answered at all. Respondents to the online survey were more avid than the actual population and appeared to be more disgruntled than the actual population, which might be a result of disgruntled people being more motivated to respond to the online survey (non-response bias).

Finally, an earlier random telephone survey allowed for comparisons between 2006 and 2008 samples (Responsive Management, 2009b). This comparison indicated a 2% decrease in saltwater fishing participation between the two study years ($\chi^2 = 7.08$, $df = 2$, $p < .05$). Because the samples for both telephone surveys were generated based on probability sampling, the two studies could be compared for statistical differences. An advantage of scientifically valid surveys is the ability to measure trends over time.

Conclusion

Online surveys are frequently not representative of the population of interest, can yield biased results, and may lead to invalid conclusions. With the exception of closed populations surveyed with specific safeguards in place (e.g., limiting the number of times a person can complete the survey), online surveys should be viewed cautiously. Obtaining representative, unbiased, scientifically valid results from online surveys is not possible at this time. There is no such thing as a complete and valid e-mail sample for most populations. Some individuals are systematically excluded (i.e., the definition of bias), and there is limited control over who completes the survey or how many times they complete the survey. These biases increase in a step-wise manner, starting out with 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.

Those not trained in survey research might assume that 10,000 responses are more accurate than 1,000. As shown by the North Carolina Sunday hunting study discussed earlier, however, earlier, *it is the method of sampling, not the sheer number of responses,* that determines sample validity. No matter how many responses are obtained, if the sample is not selected and managed properly by the researcher, the results cannot be extrapolated to the larger population. With careful and scientifically valid sample selection, a small sample size can be representative of the total population under study.

In the context of fish and wildlife management, scientifically valid human dimensions research is essential to making sound decisions. This article demonstrates the need for human dimensions research that is based on sound methodology, not on guesswork or exclusion of entire segments of a population. Just as fish and wildlife professionals rely on scientifically valid, unbiased research methods to manage habitat and species, management of the human element of fish and wildlife should be approached with equal rigor (Duda, 1986; Vaske, 2008).

References

American Association of Public Opinion Researchers. (2009). *Bad samples*. Retrieved April 3, 2009, from http://www.aapor.org/badandworsesamples

Bethlehem, J. (2008). *How accurate are self-selection web surveys?* The Hague/Heerlen: Statistics Netherlands.

Couper, M. P. (2000). Web surveys: A review of issues and approaches. *Public Opinion Quarterly, 64*, 464–494.

Dillman, D. A. (2000). *Mail and Internet surveys*. New York: Wiley.

Dillman, D. A., Smith, J. D., & Christian, L. M. (2009). *Internet, mail, and mixed-mode surveys: The tailored design method*. Hoboken, NJ: Wiley.

Duda, M. D. (1986). Wildlife management: The human element. *Florida Wildlife*, *39*(5), 15–17.

Responsive Management. (2006a). *Arizona big game hunt permit tag draw study.* Study conducted for the Arizona Game and Fish Department. Harrisonburg, VA: Author.

Responsive Management. (2006b). *Coastal resource management customer survey.* Study conducted for the National Oceanic and Atmospheric Administration Coastal Services Center. Harrisonburg, VA: Author.

Responsive Management. (2006c). *Sunday hunting in North Carolina: Results of a telephone survey of North Carolina residents.* Study conducted for the North Carolina Wildlife Resources Commission. Harrisonburg, VA: Author.

Responsive Management. (2007). *Market inventory and needs assessment.* Study conducted for the Georgia Coastal Management Program and the Sapelo Island National Estuarine Research Reserve. Harrisonburg, VA: Author.

Responsive Management. (2009a). *Comparison of a telephone survey and a Web-based survey regarding shellfishing in South Carolina.* Study conducted for the South Carolina Department of Natural Resources. Harrisonburg, VA: Author.

Responsive Management. (2009b). *South Carolina saltwater anglers' participation in and satisfaction with saltwater fishing and opinions on saltwater fisheries management.* Study conducted for the South Carolina Department of Natural Resources. Harrisonburg, VA: Author.

Siceloff, J. (2008). *About the Sarah Palin poll*. Retrieved April 3, 2009, from http://www.pbs.org/now/palin-poll.html

Vaske, J. J. (2008). *Survey research and analysis: Application in parks, recreation and human dimensions*. State College, PA: Venture Publishing.

Virginia Polytechnic and State University/Responsive Management. (2006). *An assessment of public and hunter opinions and the costs and benefits to North Carolina of hunting on Sunday*. Report prepared for the North Carolina Wildlife Resources Commission.

Zhang, Y. (1999). Using the Internet for survey research: A case study. *Journal of the American Society for Information Science 51*(1), 57–68.