Journal of Extension
www.joe.org

June 2010
Volume 48 Number 3
Article Number 3FEA2

Return to Current Issue

Assessing Awareness of Water Quality: Comparing Convenience and Random Samples

Linda Stalker Prokopy
Assistant Professor
Department of Forestry and Natural Resources
Purdue University
West Lafayette, Indiana
lprokopy@purdue.edu

Alicia Molloy
Watershed Planning Specialist
Department of Forestry and Natural Resources
Purdue University
West Lafayette, Indiana
abhaley@purdue.edu

Aaron Thompson
Ph.D. Candidate
Department of Forestry and Natural Resources
Purdue University
West Lafayette, Indiana
awthomps@purdue.edu

Don Emmert
Water Quality Educator
Tippecanoe County Soil and Water Conservation District
Lafayette, Indiana
don.emmert@in.nacdnet.net

Abstract: It is increasingly recognized that knowledge of the target audience can strengthen education and outreach programs. The study described here explores whether the method of collecting data from a target audience influences results. Using data for the same questions collected at a county fair, through the Internet, in a mail survey, and in a phone survey, our findings indicate that data collected through convenience surveys are not representative of the larger population. An implication of this finding is that organizations conducting surveys need to clearly define their goals for the survey data before choosing a method of survey delivery.

Introduction

Surveys about awareness of water quality issues are increasingly being conducted, but there is no empirical evidence that one method of collecting data is superior to another. Surveys about a community's awareness towards water quality can help generate important information that can inform the planning and evaluation stages of watershed projects (Prokopy et al., 2009). Surveys can help groups focus scarce funds to make the biggest impact. Understanding what different types of people within a community know can help refine education and outreach messages (USEPA 2002, 2003; Coburn & Donaldson, 1997). It is understood that target audiences need to be clearly identified to ensure maximum benefits from education programs (Shepherd, 1999). Collecting social data at two points in time can also provide evidence that outreach messages are effective (USEPA, 2005), and collecting baseline data is important for successful evaluation of an educational program (Shepherd, 2002).

Surveys about water quality are conducted by watershed groups, Extension personnel, and researchers across the United States. These surveys are commonly conducted at public meeting places such as county fairs. For example, a survey of Soil and Water Conservation Districts (SWCDs) that we conducted in Indiana revealed that they frequently collect data about water quality awareness at venues like county fairs. Surveys are also conducted through more rigorous methods, including mail and phone surveys (e.g., Blaine & Smith, 2007; Larson, 2009). This article compares data about water quality awareness collected through four different methods to determine whether there is a significant difference in results based upon method.

Surveys handed out at county fairs are referred to as "convenience surveys" in the literature. Without a randomly drawn sample, where each member of a population has an equal opportunity to be selected to participate, it is unclear if convenience surveys truly represent awareness in the community as a whole. Guidance texts for research design all caution against non-random sampling because it is impossible to generalize to the larger population (e.g., Chambliss & Schutt, 2006; Walliman, 2005). In other research that has compared convenience surveys to surveys conducted with a random sample, results have been mixed. Luschei, Hammond, Boerboom, and Nowak (2009) found no substantive differences between a convenience sample of on-farm research cooperators and a random sample of farmers in Wisconsin. In other research areas, however, significant differences have been found between convenience and random samples (Arnett & Rikli, 1981; Hultsch, MacDonald, Hunter, Maitland, & Dixon, 2002).

Given the prevalence of convenience surveys, it is essential to determine whether these surveys produce reliable data. To the best of our knowledge, no one has compared the results of water quality awareness surveys from convenience and random samples. Given increased calls for the collection of social data related to water quality (Prokopy et al. 2009; USEPA 2002, 2003), it is time to assess the best ways to collect this type of information. Clearly many communities, SWCDs, watershed groups, Extension offices, and others involved in water quality are cash-strapped and looking for the most cost-effective means of collecting quality data. Convenience samples are without a doubt cheaper and easier to conduct than random samples. Conducting surveys using "correct" methods costs more money, takes more time, and requires more attention.

Materials and Methods

We compare four survey methods in this article: survey handed out at a county fair (convenience sample), Web survey (convenience sample), mail survey (random sample), and phone survey (random sample). These different approaches are compared in Table 1 and discussed below.

All the surveys were conducted in Tippecanoe County, Indiana. Tippecanoe County is located in

North-Central Indiana in the Midwestern Heartland. It is the home of Purdue University and contains the Lafayette-West Lafayette metropolitan area. It is, however, a largely rural county. Only 12% of the land use is urban, while 73% is farmland (USDA 2002).

A 12-question survey was handed out to visitors at a water quality booth at the county 4-H fair in the summer of 2006. The survey was titled "What's in Your Stormwater" and was the result of a collaboration between the Tippecanoe County Soil and Water Conservation District and the Tippecanoe County Partnership for Water Quality. The fair has a large vendor and education tent where visitors can pick up information and learn more about opportunities in the county. Two hundred fifty-two surveys were completed, but it is unknown how many people saw the booth and chose not to complete the survey.

This same survey was also posted on various Web sites, including county government and Tippecanoe Municipal Separate Sewer and Stormwater System (MS4) entity sites throughout 2006. Information about the survey was distributed through informational door hangers, word-of-mouth, and university classes. Participants could voluntarily include their name as they completed the survey; a quick scan of the completed surveys shows a number of university professors completed the survey. One hundred ninety surveys were completed, but, as with the county fair survey, it is impossible to know how many people learned about the survey and chose not to take it.

An 8-page survey was mailed to a randomly selected group of 1,000 households in Tippecanoe County in the fall of 2006. The addresses were purchased from Survey Sampling International. The survey was mailed following Dillman's Tailored Design Method (2000). The survey contained questions about a variety of topics related to the Wabash River. It included seven of the same questions (worded exactly the same way) that were included in the county fair survey and the Web survey. We do not think bias from other survey questions was an issue because the questions we are examining here are fact-based questions. There were 346 respondents and 72 undeliverable surveys, leading to a response rate of 37.3%.

Finally, we conducted a phone survey of people who did not return the mail survey in Spring 2008. Using phone listings, we gathered phone numbers for people who did not respond to the mail survey. There were a total of 142 people we could not find phone numbers for (24% of the total non-responders). We then randomly sorted the list with phone numbers and called these people until 50 were reached. Twenty-seven people refused to do the survey, leading to a cooperation rate of 65%. This survey contained only the 12 questions from the county fair and Web survey.

Table 1.
Summary of Data Collected

Survey	Survey Type	Date	Number of Respondents	Method
Booth at county 4-H fair	Convenience	Summer 2006	252 (can't calculate response rate)	Visitors to water quality booth were asked to fill out a 12 question survey.
Web survey	Convenience	Throughout 2006	190 (can't calculate response rate)	Web link to 12 questions survey was distributed through door hangers, word-of-mouth, and university classes.

Mail survey	Random	Fall 2006	346 (response rate 37%)	Survey (including 8 of 12 questions) mailed to randomly selected residents of county.
Phone survey	Random	Spring 2008	50 (cooperation rate 65%)	People who did not respond to mail survey were called and asked to respond to 12 questions included on earlier surveys.

Based on the research methodology literature that suggests that convenience samples will not provide generalizable information, our a priori hypotheses were the following.

1. There will be significant differences in awareness of water quality issues across the survey modes.

2. People who voluntarily responded to the Web survey that was dropped at doors only once with no personal contact will have more personal interest in the topic, and subsequently more awareness, than people who filled out the survey at the fair.

3. The phone and mail survey respondents (random surveys) will have lower levels of awareness than both the county fair and Web respondents (convenience surveys).

4. People who did not respond to the mail survey (i.e., people contacted by phone) will have lower levels of awareness than mail respondents as they are assumed to have the least interest in the topic area.

These hypotheses were tested in SPSS 16.0. Hypothesis 1 was tested using a one-way between groups Analysis of Variance (ANOVA) test. The independent variables for this test are the survey mode used and the dependent variables are the survey questions. Hypotheses 2-4 were tested with difference of means t-tests in which the independent variables are again the survey mode used and the dependent variables are the survey questions.

Note that these surveys were not all conducted at the same time; however, no unusual efforts were made towards increasing awareness about water quality issues in the community during this time. The community has continued to increase the numbers of storm drain markers, but they were prevalent even before the surveys began. The community recently received a watershed-planning grant, but no outreach for this grant was done before these surveys were conducted. There are ongoing educational efforts in schools by the SWCD, but we have no reason to expect that the "true" levels of awareness in the community have changed during this time period.

Results and Discussion

Table 2 summarizes the results based upon percentage of correct and incorrect responses for the 10 questions for which a "correct" answer could be determined. A quick glance through these questions show that

responses are somewhat similar for Web and fair surveys and somewhat similar for mail and phone surveys. However responses are dissimilar between the convenience and random surveys. As expected, Web survey responders appear to have the most knowledge; they are also most likely to have seen, read, or heard about things they can do to reduce stormwater pollution during the year. Sixty-six percent of Web responders know they live in a watershed, compared to 52% of fair responders and only 20% of phone responders. Almost twice as many Web and fair responders know the best way to wash a car compared to phone and mail responders.

Table 2 also shows significance levels resulting from the between groups ANOVA test. At a significance level of .05, there is a significant difference between groups for over half the questions. This suggests that the multiple sources of information are not providing us with equivalent data and our first hypothesis is confirmed.

Table 2.
Answers to Survey Questions Presented as Percentage Correct/Incorrect and ANOVA Significance Levels

	Convenience Surveys				**Random Surveys**				**ANOVA Significance Level**
	Web		Fair		Mail		Phone		
	Correct	Incorrect	Correct	Incorrect	Correct	Incorrect	Correct	Incorrect	
I live in a watershed	66	34	52	48	N/A		20	80	.000***
The greatest pollutant by volume in our creeks, rivers, and lakes is silt or sediment	66	34	65	35	N/A		48	52	.408
It is best to use a commercial car wash to clean your vehicle instead of washing it at home	67	33	60	40	35	64	34	66	.000***
Paint, solvents, or used motor oil may be disposed of safely into a storm sewer	97	3	94	6	96	4	88	12	.116
	70	30	61	39	N/A		46	54	.001***

Muddy washwater may be disposed of safely into a storm sewer									
Grass clippings, garden trimmings, and fallen leaves can be a source of water pollution	82	18	79	21	58	42	62	38	.000***
A properly functioning septic system should be pumped, cleaned and inspected every 3 to 5 years	75	25	78	22	63	37	82	18	.056*
Most street and parking lot drains flow directly into a nearby creek, river, or lake	64	36	65	35	61	39	46	54	.074*
Pet waste should be picked up only as a courtesy to others, not because it causes water pollution	85	15	82	18	51	49	68	32	.000***
All fertilizers and chemicals	98	2	92	8	87	13	86	14	.001***

applied to lawns stay within the yard.									
* significant at .10 level; ** significant at .05 level; *** significant at .01 level									

The difference of means t-tests used to examine hypotheses 2-4 are included in Table 3. Exploring the results for the second hypothesis shows that there are significant differences at the 0.05 level between Web and fair surveys for only two questions. We therefore conclude that in general there is not an overwhelmingly significant difference in the two modes of delivering convenience surveys.

To compare the differences between convenience surveys and random surveys (hypothesis 3), we created two groups in the dataset: the data for the fair and Web survey were combined into one group, and the data for the phone and mail survey were combined into another group. Comparing the means between these two new groups revealed significant differences (at the 0.05 level) in correct answers for all but two variables. The two variables that do not differ significantly relate to disposal of paints and solvents and where storm drains flow to (this latter variable is significant at 0.10).

We conclude that the choice of a convenience versus a random survey is an important one. In all cases, the mean awareness level was higher for respondents taking a convenience survey, suggesting that these types of surveys are likely over-estimating the level of awareness than is actually present in the population.

Table 3.
Difference of Means Test Results for Hypotheses 2, 3 and 4

	Hypothesis 2: Web Surveys vs. County Fair Surveys		**Hypothesis 3: Convenience Surveys vs. Random Surveys**		**Hypothesis 4: Mail Surveys vs. Phone Surveys**	
	Mean Difference	Significance	Mean Difference	Significance	Mean Difference	Significance
I live in a watershed	14	.006***	38†	.000***	N/A	
The greatest pollutant by volume in our creeks, rivers, and lakes is silt or sediment	1	.989	17†	.015**	N/A	
It is best to use a commercial car wash to clean your	7	.173	27	.000***	2	.831

vehicle instead of washing it at home						
Paint, solvents, or used motor oil may be disposed of safely into a storm sewer	3	.175	0	.926	8	.010***
Muddy washwater may be disposed of safely into a storm sewer	8	.083*	19†	.010***	N/A	
Grass clippings, garden trimmings, and fallen leaves can be a source of water pollution	2	.538	22	.000***	4	.567
A properly functioning septic system should be pumped, cleaned and inspected every 3 to 5 years	0	.490	11	.000***	19	.008***
Most street and parking lot drains flow directly into a nearby creek, river, or lake	2	.724	6	.079*	15	.050**
Pet waste should be	3	.350	30	.000***	17	.024**

picked up only as a courtesy to others, not because it causes water pollution						
All fertilizers and chemicals applied to lawns stay within the yard.	6	.011**	8	.000***	1	.858

† This comparison compares Web and fair surveys (convenience surveys) to phone surveys (random surveys). Note that mail surveys are not included in the comparison for this question because it was not asked on the mail survey.
* significant at .10 level; ** significant at .05 level; *** significant at .01 level

Examining relationships specifically between phone and mail responders (hypothesis 4) shows that there are four significant differences at the 0.05 level. In half of these cases, phone responders were less aware of water quality issues than the mail responders. However, for the question about septic systems and the question about pet waste, phone responders were more aware than mail responders. With this disparate information, we cannot concretely answer our hypothesis.

Clearly, there were limitations in our studyâ† even the phone survey only reached people with landlines. Close to 25% of our non-mail responders did not have phone numbers. Some of these people likely moved between the mail and phone survey; however, a significant percentage of them just didn't have listed phone numbers. Given the increasing numbers of people without landlines, this method of verifying the reliability of mail survey data is not as immune from criticism as it once was. However, there is not yet a better alternative.

Summary and Conclusions

As hypothesized, for the most part, respondents to the convenience surveys had higher levels of awareness about water quality than respondents to the random surveys. However, it is important to consider context. For example, Luschei et al. (2009) found that a convenience survey did provide reliable data in Wisconsin. Their study looked at on-farm research, and they compared voluntary cooperators to a larger sample of farmers and surveyed both sets about farmer behaviors and attitudes related to weed management.

Organizations and Extension offices conducting surveys should clearly identify their goals before choosing a method of survey delivery. These goals need to be considered in light of available resources. Many organizations do not have the funds to conduct a random survey such as the mail survey described here. To that end, if the purpose of a survey is to raise awareness, and the data are not actually going to be used, then perhaps there is nothing wrong with the public venue approach that is frequently used. This approach has the additional benefit of allowing for interaction with the community, whereas phone calls and mailings are less personal.

However, the results of these surveys are sometimes used to inform education and outreach. Based upon our findings, this is potentially problematic. If education and outreach campaigns are going to be based upon survey data, organizations should think very seriously about using a random sampling method. If data are not collected using a random sample, then resources may not be directed in the most efficient way. Watershed projects increasingly focus on diverse audiences and need a method that can capture diversity in awareness across the entire population, not just people who attend a 4-H fair.

If organizations are collecting survey data for evaluation purposes using a longitudinal approach with baseline and follow-up data, then it depends on what the organizations want to know. Do they want to know if the population attending the state fair has changed? Or do they want to know that a broader population has changed? These are important questions to answer before conducting the baseline survey.

Future studies could examine the impacts of a convenience sample on questions that go beyond simple awareness. For example, Prokopy et al. (2009) suggest that questions related to attitudes, behaviors, and constraints are also important for planning and evaluating watershed management efforts. However, there is no reason to expect the findings of such a study to differ from those presented here.

Acknowledgments

Thanks to all the volunteers who helped distribute the survey at the 4-H fair and the link to the Web survey. Thanks to Erin Barnes, Katie Crosley, and Sarah Bechtol for conducting the phone surveys and Cindy Salazar for supervising their work. The mail survey was conducted as part of the Living Laboratories on the Wabash project and funded by a grant from the Center for the Environment at Purdue University.

References

Arnett, B., & Rikli, R. (1981). Effects of method of subject selection (volunteer vs. random) and treatment variable on motor performance. *Research Quarterly for Exercise and Sport* , 52 (4),433-440.

Blaine, T. W.. & Smith, T. (2007). From water quality to riparian corridors: Assessing willingness to pay for conservation easements using the contingent valuation method. *Journal of Extension* [On-line], 44(2) Article 2FEA7. Available at: http://www.joe.org/joe/2006april/a7.php

Chambliss, D. F., & Schutt, R. K. (2006). *Making sense of the social world: Methods of investigation*. Thousand Oaks: Pine Forge Press.

Coburn, J., & Donaldson, S. (1997). Reaching a new audience. *Journal of Extension* [On-line], 35(1) Article 1FEA3. Available at: http://www.joe.org/joe/1997february/a3.php

Hultsch, D. F., MacDonald, S. W. S., Hunter M. A., Maitland S. B., & Dixon R. A. (2002). Sampling and generalisability in developmental research: Comparison of random and convenience samples of older adults. *International Journal of Behavioral Development,* 26 (4).

Larson, K. (2009). Social acceptability of water resource management: A conceptual approach and empirical findings from Portland, Oregon. *Journal of the American Water Resources Association*, 45(4): 879-893.

Luschei, E. C., Hammond C. M., Boerboom, C. M., & Nowak, P. (2009). Convenience sample of on-farm research cooperators representative of Wisconsin farmers. *Weed Technology,* 29: 300-307.

Prokopy, L. S, Genskow, K., Asher, J., Baumgart-Getz, A., Bonnell, J. E., Broussard, S., Curtis, C., Floress, K., McDermaid, K., Power, R., & Wood, D. (2009). Designing a regional system of social indicators to evaluate nonpoint

source water projects. *Journal of Extension* [On-line], 47(2) Article 2FEA1. Available at: http://www.joe.org/joe/2009april/a1.php

Shepherd, R. (1999). Making our nonpoint source pollution education programs effective. *Journal of Extension* [On-line], 37(5) Article 5FEA2. Available at: http://www.joe.org/joe/1999october/a2.php

Shepherd, R. (2002). Evaluating Extension-based water resource outreach programs: Are we meeting the challenge?. *Journal of Extension* [On-line], 40(1) Article 1FEA3. Available at: http://www.joe.org/joe/2002february/a3.php

US Department of Agriculture. (2002). *Census of Agriculture*. Retrieved November, 2005 from: http://www.nass.usda.gov/Census_of_Agriculture/index.asp

U.S. Environmental Protection Agency (USEPA). (2002). *Community culture and the environment: A guide to understanding a sense of place*. USEPA Washington, D.C.

U.S. Environmental Protection Agency (USEPA). (2003). *Getting in step: A guide for conducting watershed outreach campaigns*. USEPA Washington, D.C.

U.S. Environmental Protection Agency (USEPA). (2005). *Stormwater Phase II final rule: Public education and outreach minimum control measure*. Fact Sheet 2.3. EPA 833-FOO-005. USEPA Washington, D.C.

Walliman, N. (2005). *Your research project: A step-by-step guide for the first-time researcher*. Second ed. Thousand Oaks: Sage Publications.
