Recommended Study Design for a Survey to Evaluate the Effectiveness of Mississippi Delta Fish Advisories

The full report can be found on our website at:
http://www.epa.gov/fishadvisories/

Final Report

June 29, 2007

Prepared for

The Office of Water, Office of Science and Technology
U.S. Environmental Protection Agency
1200 Pennsylvania Avenue, NW
Washington, DC 20460

3 Recommended Survey Methodology

This section presents the recommended survey methodology, including the survey mode, the survey universe, the sample design, and survey procedures.

3.1 SURVEY MODE

The workgroup evaluated the advantages and disadvantages of alternative survey modes and decided that a combination of on-the-bank interviews and a household survey would be the best approach for reaching the target population.

The target population for the survey includes

- sport and subsistence fishers who harvest noncommercial fish from the Mississippi Delta advisory area, and
- individuals in the Mississippi Delta area who consume wild-caught fish from the advisory area.

The workgroup evaluated the advantages and disadvantages of alternative survey modes (see Table 3-1). The workgroup did not consider using a mail survey because of concerns about illiteracy among some members of the target population. The workgroup decided against a telephone survey because in some counties, more than 25% of the population does not have a telephone. Also, the workgroup believed that the target population would not be very receptive to a telephone survey.

The workgroup decided that a combination of on-the-bank interviews and a household survey would be the best approach to reach the target population. The recommended survey approach combines a convenience sample for the on-the-bank survey with a probability-based sample in four counties for the household survey.

Table 3-1. Advantages and Disadvantages of Alternative Survey Modes

	Random-Digit-Dial (RDD) Telephone Survey	Household Survey	On-the-Bank Survey
Expected response rate	Low, less likely to get "buy-in" for survey over the phone	High[a]	High[b]
Interviewers	Unlikely that "locals" would be used to conduct the survey	Ability to use "locals" to conduct the survey	Ability to use "locals" to conduct the survey
Screening costs	High, will have to screen to identify fishers or people who consume fish[c]	High, will have to screen to identify fishers or people who consume fish	None
Representativeness	Would exclude households without telephone numbers; 75%–95% of households in Delta have telephones	Could design survey to be representative of the area sampled	Difficult and costly to design statistically representative survey; would exclude individuals who fish from small private ponds; individuals without a fishing license may be reluctant to participate if approached on the bank
Questionnaire length	To maximize response rate, interview length should be 10 to 15 minutes	30 to 45 minutes	15 to 20 minutes
Questionnaire type	Limited to closed-ended questions, limited ability to probe	Could ask some open-ended questions, easier for interviewer to probe and clarify questions	Could ask some open-ended questions, easier for interviewer to probe and clarify questions
Relative cost	Low	High	High

[a] Brown, Xu, and Toth (1998) achieved an 88% response rate for a household survey conducted in two Delta communities.

[b] Burger and other researchers have typically achieved response rates of greater than 90% for on-the-bank surveys (Burger and Waishwell, 2001; Pflugh et al., 1999; and Campbell et al., 2002).

[c] In a household survey of two Delta communities, Brown, Xu, and Toth (1998) found that 37% of households fished.

3.2 SURVEY UNIVERSE

The workgroup decided that designing a survey approach that would be representative of all counties in the Mississippi Delta advisory area was not necessary. Instead, the workgroup purposively selected four counties for the survey. Counties were selected to include a mix of rural and nonrural areas and areas with major water resources affected by the advisory

(e.g., Roebuck Lake, Moon Lake, Lake Washington, and Bee Lake). The selected counties are Coahoma, Holmes, Leflore, and Washington. Figure 3-1 illustrates the location of the four counties within the advisory area. Only the part of Holmes County that is within the advisory area will be included in the survey. Table 3-2 provides sociodemographic information on the four counties from the most recently available Census data.

Figure 3-1. Location of the Four Counties within the Advisory Area

Table 3-2. Characteristics of the Four Counties for the Survey

County	Major Cities	Population, 2006	Number of Housing Units, 2005	Land Area, 2000 (sq mi)	Persons per Square Mile, 2000	Race			Median Household Income, 2004
						White	Black	Hispanic	
Coahoma	Clarksdale, Jonestown	28,420	11,587	554	55	26%	73%	1%	$23,560
Holmes	Durant, Lexington	20,866	8,725	756	29	19%	81%	1%	$20,295
Leflore	Greenwood	35,752	14,612	592	64	29%	71%	2%	$22,709
Washington	Greenville	58,007	24,958	724	87	32%	67%	1%	$25,455

Source: U.S. Census Bureau. State and County Quick Facts. http://quickfacts.census.gov/qfd/states/00000.html.

3.3 SAMPLE DESIGN

The recommended sample design developed by RTI for the on-the-bank and the household surveys is described below.

3.3.1 On-the-Bank Survey

The primary advantage of an on-the-bank survey is that it greatly increases the likelihood of reaching individuals fishing in the advisory area who possibly consume the contaminated species. Because of the difficulty and complexity of designing a probability-based sample for an on-the-bank survey of multiple water resources that are geographically dispersed, the workgroup recommended using a convenience sample for the survey.

For the on-the-bank survey, RTI recommends interviewing approximately 400 fishers allocated to each county in proportion to their total population for the advisory area (Table 3-3), with interviews conducted at four to five different water bodies in each county.[1] Table 3-4 identifies recommended water bodies for the on-the-bank survey. Appendix D provides maps of the counties, including the location of many of the recommended water bodies.

3.3.2 Household Survey

For the household survey, RTI recommends using a stratified equal-probability two-stage design. With this design, census tracts are selected in the first stage, and households within the selected tracts are selected in the second stage. The design will resort to probability-proportional-to-size sampling, which means that first-stage units are selected with probabilities proportional to some measure of size. For example, if the measure of size is total households, a tract with 100 households will have twice the probability of selection as a tract with 50 households. However, in the second stage a fixed number of households will be selected, leading to overall equal probabilities of selection and evenly distributed interviewer workload over the selected census tracts. Because the recommended design is a probability-based design, inferences

[1] For Holmes County, only the portion of the county in the advisory area was used for determining the sample allocation.

Table 3-3. Recommended Sample Allocation for the On-the-Bank and Household Surveys

County	On-the-Bank Survey	Household Survey
Coahoma	80	195
Holmes	55	140
Leflore	100	250
Washington	165	415
Total	400	1,000

Table 3-4. Recommended Water Bodies for the On-the-Bank Survey

County	Water Bodies	Comments
Coahoma	Moon Lake	Major lake
	Long Lake	
	Mile Marker 98 Pond	
	Sunflower River	Access points are Johnson's low water crossing, Hopson, and Clarksdale
	DeSota Lake	
Holmes	Bee Lake	Private lake, will need to request permission
	Horseshoe Lake	
	Jonestown Cutoff	
	Bear Lake	Located in Hillside National Wildlife Refuge
	Little Legal Lake	Located in Humphreys County, on the county line
	Other	The Department of Wildlife and Fisheries can provide driving route that covers 10 small lakes and creeks
Leflore	Roebuck Lake	Advisory recommends not eating any buffalo fish from this lake
	Little Round Lake	
	Matthews Break	Located in National Wildlife Refuge
	Blue Lake	
	Yazoo River	The boat ramp at Greenwood also provides access to Yalobusha River and Tallahatchie River
	Sixmile Lake	
Washington	Washington Lake	Major lake
	Black Bayou	Located at Leroy Percy State Park
	Deer River	Located in Hollandale
	Big Sunflower River	Access points are at Route 12 and Murphy Landing
	Lake Jackson	Boats only

can be made to the population of all occupied housing units in the four-country survey area. Additionally, the design will allow for comparisons among key sociodemographic variables such as gender and race.

The sample size, first and second stage selection procedures, stratification, and weighting procedures for the household survey are described below.

Sample Size

The universe for the survey consists of all occupied housing units in the four counties of Coahoma, Leflore, Holmes (advisory area only), and Washington.

RTI recommends a sample size of 1,000 completed interviews, allocated to each county in proportion to the total population for the advisory area (Table 3-4). The combined sample of 1,000 will result in a 95% confidence interval (CI) of (0.46, 0.54) for a proportion estimate of 0.50, assuming a design effect of 2. This level of precision will also result in considerable power for identifying differences between estimates. Table 3-5 provides corresponding CIs for other sample sizes. For example, if the number of completed interviews is decreased to 500, then the 95% CI would be (0.44, 0.56).

Table 3-5. Corresponding 95% CIs for Varying Sample Sizes for a Proportion Estimate of 0.50, Assuming a Design Effect of 2

Sample Size	Lower-Bound CI	Upper-Bound CI
250	41%	59%
500	44%	56%
750	45%	55%
1,000	46%	54%

First Selection Stage

In the first stage of selection, a total of 50 primary sampling units (PSUs), defined as tracts (i.e., units that contain approximately 500 households) will be sampled from the four counties. These will be sampled with probabilities proportional to size, where size is an estimate of the total population or total number of households within each county for the advisory area.

Stratification

PSUs will be stratified by geography, access to water bodies of interest, and various demographic and socioeconomic characteristics to ensure a good spread of the sample across all segments of the population. Implicit stratification will be implemented by sorting all PSUs by the stratification variables and then drawing a systematic sample based on a random start and a fixed sampling interval, with the measure of size for each PSU being used as the sampling weight for that PSU.

Second Selection Stage

Once the 50 PSUs have been selected, a listing of all addresses in the selected PSUs will be made. Commercial lists are available through sources such as Marketing Systems Group and Survey Sampling, Inc. The list of all addresses in the PSUs is the sampling frame for the second selection stage.

The number of households to sample in each selected PSU will be 20. The households in each PSU will be sorted in geographical sequence and a sample taken, after a random start, of every nth household where n is the result of dividing the total number of households in the PSU by 20. A simplified version of the "half-open-interval" (HOI) is recommended to account for existing households that are not included on the list of addresses. With HOI, the interviewer is given the sampled address and the next address on the list. The interviewer is instructed to ensure that there are no housing units between the two addresses. If there are, a simple algorithm can be used to help decide whether the found housing unit should be interviewed.

Weighting

Although there are unequal probabilities of selection at each stage, the overall two-stage probabilities of selection are equal:

$$p = (50 * Mos_i)/\sum Mos_i * (20/Mos_i) = 1{,}000 / \sum Mos_i$$

where

p = the overall probability of selection, and

Mos = the measure of size for PSU_i (e.g., the total number of households in PSU_i).

However, the measures of size in the first stage are based on best available data, whereas in the second stage they are

based on the number of addresses available. To the extent that the two differ, it will be necessary to make adjustments to the second stage sample size or use weighting adjustments. For weighting purposes, adjustments will need to be made for other deviations from the original design, for variable nonresponse, and for final poststratification adjustment to known population benchmarks. RTI recommends using SUDAAN, a statistical analysis software tool, to calculate estimates incorporating the complexity of the design, namely, the stratification and clustering. Using the recommended design, the design effect should not exceed 2 for most estimates.

3.4 SURVEY PROCEDURES

The recommended survey procedures for the on-the-bank and household surveys are described below. Also included is a discussion of the selection and training of interviewers, the selection of individuals to interview, the use of incentives, the time period for field data collection, and other survey procedures.

3.4.1 Selection and Training of Interviewers

The workgroup recommended using individuals who are local to the area to conduct the on-the-bank and household surveys. Using “locals” to conduct the interviews will facilitate access to the interview area and thus help maximize the response rate for the surveys. Additionally, these individuals will be familiar with the Delta culture and water resources. For example, Brown, Xu, and Toth (1998) used local school teachers to conduct a survey in two Delta communities and achieved an 88% response rate. The survey contractor should work with a local university or college such as Delta State University in Cleveland, Mississippi, or Mississippi State University in Starkville, Mississippi, to facilitate identifying and training qualified individuals for conducting the interviews. Additionally, it is important to match interviewers to the area or water resource being surveyed based on race (e.g., using African American interviewers to interview individuals in predominantly African American neighborhoods). In some locations it may be useful to send two interviewers to each neighborhood and water resource—one white and one African American—and then match the interviewer to the race of the respondent.

The success of any survey depends on staff who are well trained, motivated, and skilled in collecting high-quality data. Thus, the workgroup recommended that interviewers be thoroughly trained. In addition to training on general interviewing methods and refusal/avoidance techniques, interviewers should participate in specialized training sessions on conducting on-the-bank interviews and techniques for collecting data on fish consumption. As part of the training, interviewers should conduct mock interviews to practice administering the questionnaire so that they become familiar with following the skip patterns and selecting the correct response options.

3.4.2 Selection of Individuals to Interview

The recommended procedures for selecting individuals to interview for the on-the-bank and household surveys are described below.

On-the-Bank Survey

No screening questions are necessary for the on-the-bank survey because any individual aged 18 years or older who is observed fishing is eligible for the survey. The interviewer should verify the respondent's age if the interviewer is not sure if the respondent is at least 18 years old. For individuals fishing alone, interviewers should approach each person observed fishing and request an interview. For individuals fishing in groups, interviewers should select one fisher per group to be interviewed. Interviewers should purposively select an individual to interview to maintain a mix of races and genders.

The workgroup recommended using a survey approach similar to that employed by Burger and other researchers (Burger and Waishwell, 2001; Pflugh et al., 1999; and Campbell et al., 2002). In this approach, interviewers are assigned to certain water bodies and are responsible for visiting a certain number of sites each day. Not all sites are visited each day, but all sites are visited on Sunday and Saturday to get a sample of fishers at each site on different days of the week and to determine if some days are more popular than others. This approach helps ensure the inclusion of the widest possible cross section of the fisher population in the survey. Interviewers will drive to the different sites at least twice each field day to interview fishers who might arrive at different times during the course of the day.

Household Survey

Interviewers will be provided with a list of the selected households to interview. For a selected household to be eligible for the survey, it must meet the following inclusion criteria: (1) an adult household member has been fishing in the Delta in the past year or (2) a household member has consumed wild-caught fish from the Delta in the past year (see Questions E.1 and E.2 of the questionnaire). If eligible, the interviewer then selects an individual within the household to interview. For households that fish, interviewers should select the adult who fishes most often; for households that consume fish, interviewers should select the adult who prepares most of the wild-caught fish consumed by household members. For households that fish *and* consume fish, interviewers should randomly select the person to interview. In all cases, a follow-up appointment should be scheduled if the selected individual is not available.

3.4.3 Use of Incentives

The workgroup did not recommend providing respondents with monetary incentives. Other researchers have achieved high response rates for similar surveys and populations without using monetary incentives (Burger and Waishwell, 2001; Pflugh et al., 1999; Campbell et al., 2002; Brown, Xu, and Toth, 1998). To encourage response, interviewers should inform respondents that the survey results will be distributed locally (e.g., posted on the MDEQ Web site or in county government offices).

3.4.4 Time Period for Field Data Collection

The workgroup recommended conducting the survey from April to June 2008. Most fishing by commercial fishers is conducted between March and July (Brown and Toth, 2001); thus, this time period will capture sport and subsistence fishers.

3.4.5 Other Survey Procedures

Interviewers should begin the interview with a conversation about fishing to “break the ice” and to help establish rapport with the respondent. Additionally, interviewers should not overdress (e.g., men should not wear ties) so that respondents are at ease with the interviewer.

Interviewers should use visuals to facilitate respondent understanding of certain questions. As identified in the

questionnaire, interviewers should use a map so that respondents understand the area affected by the advisory and the location of Roebuck Lake (see Questions E.1 and 48). Also, fish models should be used to illustrate meal size for the question on fish consumption (see Question 9). After awareness of the advisory has been determined, interviewers should show respondents the sign that illustrates the fish species included in the advisory (see Questions 17 and 37). Also, interviewers should be prepared to use their hands (spread apart by about 2 feet) to show the length of a catfish greater than 22 inches.

The Hispanic population is very small in the selected counties (between 1% and 2%); thus, it is not necessary to translate the questionnaire into Spanish and conduct interviews in Spanish.