# Evaluating Measurement Error in the MRIP Fishing Effort Survey

NOAA Fisheries Service, Office of Science and Technology

May, 2023

**Executive Summary**

Since implementing the Fishing Effort Survey (FES) in 2015, we have administered several pilot studies and data analyses to evaluate non-sampling errors in the survey design. The goal of these studies is to quantify the magnitude of bias resulting from non-sampling errors, including nonresponse, non-coverage and measurement errors, and develop revised or new methods to reduce or mitigate bias. This document describes results from two studies focused on measurement error, which is the difference between a measured or reported value and the true value. Measurement error includes recall error, which can result in under-reporting (omission error) or over-reporting (telescoping error) and can often be reduced by carefully articulating survey questions and instructions, varying the order of questions and the length of reference periods, and providing memory cues.

The first study, described in the next section, evaluated recall error in the FES by comparing FES estimates, which are based upon two-month reference periods, to those from experimental questionnaires that collected data for one-month reference periods. Results from the study suggest that FES estimates are not likely to be biased from errors of omission - respondents are unlikely to forget or under-report fishing activity that occurred during two-month reference periods. Rather, respondents are more likely to over-report fishing activity when the reference period is limited to a single month - respondents telescope earlier, out-of-scope trips into the reference period. The study identifies an approach to reduce telescoping error by providing bounds, in the form of questions about additional reference periods, against which responses are based. Specifically, an experimental questionnaire that asked about two months individually produced estimates that were not significantly different from the FES, which asks about a single two-month reference period.

The second study included analytical and experimental work to evaluate recall error related to the order in which survey questions are presented. Results from this study also suggest that the predominant form of measurement error in the FES is telescoping error; respondents are more likely to report out-of-scope trips than omit trips. The current FES questionnaire design includes bounding questions (12-month shore and boat fishing) that are likely to reduce telescoping error relative to an unbounded design. However, the order in which questions are presented may not be optimal in terms of reducing telescoping error - respondents are unlikely to review the entire questionnaire, including the bounding questions, prior to reporting for the desired reference period. Consequently, the current design is likely overestimating fishing effort. A revised design that presents the 12-month fishing questions before the two-month questions would likely further reduce bias resulting from telescoping error.

# Section 1: A Comparison of Recall Error in Recreational Fisheries Surveys with One- and Two-Month Reference Periods

*North American Journal of Fisheries Management*
Published 2018. This article is a U.S. Government work and is in the public domain in the USA
ISSN: 0275-5947 print / 1548-8675 online
DOI: 10.1002/nafm.10233

ARTICLE

# A Comparison of Recall Error in Recreational Fisheries Surveys with One- and Two-Month Reference Periods

**William R. Andrews***
*National Oceanic and Atmospheric Administration Fisheries, Office of Science and Technology, 1315 East-West Highway, Silver Spring, Maryland 20190, USA*

**Katherine J. Papacostas**
*National Oceanic and Atmospheric Administration Fisheries, Office of Science and Technology, 1315 East-West Highway, Silver Spring, Maryland 20190, USA; and ECS, 2750 Prosperity Avenue, Fairfax, Virginia 22031, USA*

**John Foster**
*National Oceanic and Atmospheric Administration Fisheries, Office of Science and Technology, 1315 East-West Highway, Silver Spring, Maryland 20190, USA*

***Abstract***
**Many fisheries monitoring programs use self-administered surveys to collect data, which are subject to recall error. Recall error occurs when respondents inaccurately remember past events due to telescoping (remembering events more recently or further back in time than they occurred) or omission error (forgetting events altogether). Previous research on the effects of variable reference periods in fisheries surveys has been inconclusive due to difficulty in disentangling method effects from recall error and in determining whether estimates from shorter recall periods are less biased or more subject to telescoping. The National Marine Fisheries Service has developed a new household mail survey, the Fishing Effort Survey (FES), in which anglers are asked to recall cumulative fishing effort over the past 2 months, from which estimates of saltwater fishing effort are produced. Here, we examined how the length of the reference period may affect the FES in four U.S. states by comparing effort estimates to two feasible alternatives: (1) a survey administered monthly with both a 1- and 2-month reference period (wherein respondents were asked to recall fishing effort for each of the past 2 months individually); and (2) a survey administered monthly with a 1-month reference period. To further explore bias in the designs, we compared total effort, fishing prevalence, and mean trips per household estimates derived from the two experimental surveys. We found no significant differences between the FES and experimental survey estimates. However, we found evidence that multiple reference periods in a single survey may reduce bias for 1-month estimates. Increased understanding of (1) techniques that can reduce recall bias and (2) the trade-offs of shorter or longer reference periods will ultimately help fisheries survey designers more accurately weigh bias against survey costs and improve the quality of data used to inform management decisions.**

Self-reported data collected through retrospective recall of past events are a crucial component of a variety of social, public health, and economic research efforts (e.g., Abbott and Monsen 1979; Wright and Pescosolido 2002; Bhandari and Wagner 2006) and have been widely used to estimate recreational fishing statistics in the United States and elsewhere (e.g., Hicks et al. 1999; Ditton and Hunt 2008; Sampson 2011; Rocklin et al. 2014). Such data,

*Corresponding author: rob.andrews@noaa.gov
Received March 13, 2018; accepted September 21, 2018

however, are subject to various sources of nonsampling error, including measurement error. Memory or recall error is a type of measurement error that occurs when respondents are unable to accurately remember or recall past events (Neter and Waksberg 1964; Eisenhower et al. 2011). Recall errors are typically classified as either telescoping error or omission error (Sudman and Bradburn 1973; Chu et al. 1992). Telescoping occurs when a respondent misplaces an event in time, usually placing the event more recently in time than it actually occurred; omission error, also referred to as "recall decay," occurs when a respondent forgets an event.

Several factors are thought to affect a respondent's ability to remember and report past events, including (1) the number of events (i.e., reporting becomes more time consuming as the number of events increases); (2) the extent to which events are important or memorable (salience); (3) the frequency or regularity of events; and (4) the length of the reference period, or the time period for which recall of an activity is utilized by the respondent: longer reference periods potentially require recollection of events that are more distant as well as a greater number of events (Blair and Burton 1987). It is generally accepted that the greater the length of the reference period, the greater the expected bias due to recall error.

Identifying how to best minimize recall error while maximizing the quantity of information collected and optimizing a survey's budget remains a challenge (Clarke et al. 2008). Researchers have developed several strategies to enhance memory and subsequently reduce recall error (Sudman and Bradburn 1974). These include aided recall, which stimulates recall by providing memory cues, such as pictures or calendars; requesting that respondents consult personal records, such as bank statements or receipts; landmark procedures, which relate the reference period to a landmark event, such as a major holiday, personal milestone, or natural disaster (Loftus and Marburger 1983; Gaskell et al. 2000); adjusting the duration of the reference period (Chu et al. 1989); and bounded recall, which bounds respondent memory against a prior interview (Neter and Waksberg 1964) or a previous question within a single interview (Sudman et al. 1984). Researchers frequently utilize a combination of these approaches to improve the quality of survey responses.

Prior studies have been inconsistent with respect to the effects of reference period length on recreational fisheries survey measures (Gems et al. 1982; Chu et al. 1992; Tarrant et al. 1993; Connelly and Brown 1995, 2011; Connelly et al. 2000). For example, Gems et al. (1982) found that a 2-month reference period resulted in lower estimates of fishing activity than a 2-week reference period and attributed the difference to omission error associated with a longer reference period. In contrast, others have suggested that longer reference periods result in overestimation of fishing activity (Chu et al. 1992; Tarrant et al. 1993; Connelly and Brown 1995). Still others report no difference in reported fishing activity as a function of the duration of the reference period (Connelly and Brown 2011). An enhanced understanding of how recall affects recreational fisheries data collection programs is needed to continue improving the accuracy of recreational fisheries statistics.

One factor that may contribute to inconsistent findings is the difference among survey designs that have been utilized to examine recall error in recreational fishing surveys. For example, some studies have compared angler diaries to mail surveys with longer reference periods (e.g., Tarrant et al. 1993), while others have used mail surveys to examine one reference period and telephone surveys for another (e.g., Connelly et al. 2000). In all of these studies, the authors acknowledged that it was difficult in such designs to disentangle method effects from recall bias. Others have used the same survey methods with two different reference periods to better isolate recall bias (e.g., Connelly and Brown 2011), but they acknowledged that even in using identical methodologies, it was difficult to conclude whether shorter reference periods reduced recall error or were instead subject to more telescoping bias than longer reference periods.

The National Marine Fisheries Service (NMFS) of the National Oceanic and Atmospheric Administration has redesigned its marine recreational fisheries data collection program, creating the Marine Recreational Information Program (MRIP; see National Research Council 2006; National Academies of Sciences, Engineering, and Medicine 2017). In January 2018, the MRIP transitioned to a new survey, known as the Fishing Effort Survey (FES), to collect data about recreational shore and private boat fishing trips along the U.S. Atlantic and Gulf of Mexico coasts. The FES is a self-administered mail survey that asks household residents to report recreational saltwater fishing trips that occurred during 2-month reference periods, or "waves." These data are used to estimate fishing effort (i.e., the total number of shore and private boat fishing trips) for each of six 2-month waves as well as annual fishing effort at the end of each calendar year. The FES replaced the legacy Coastal Household Telephone Survey, a random-digit dial, landline telephone survey that NMFS had used to estimate fishing effort since 1981 (Brick et al. 2012). The FES has been identified as a more efficient and accurate approach for monitoring recreational fishing effort than the Coastal Household Telephone Survey (Andrews et al. 2014; National Academies of Sciences, Engineering, and Medicine 2017). However, the MRIP continues to examine the impacts of measurement errors, including recall error, on estimates in an effort to understand potential biases and limitations of the FES design.

Accurate statistics are essential for quantifying the effects of recreational fishing on fish stocks and developing sound, evidence-based management strategies and policies. Continuous catch and effort monitoring, for example, is needed to assess trends, evaluate the impacts of management regulations, and project how different management scenarios might influence a fishery. Minimizing biases, including recall error, in recreational fisheries surveys is therefore a necessity for effective management; large biases reduce data quality and the subsequent utility of the statistics produced from those data to fisheries scientists and managers. Understanding the magnitude of biases that occur in existing survey methods—as well as exploring methods to help mitigate such biases—can help to improve data quality so that managers are provided with the best possible scientific information to use in their decision making (National Standard 2, Magnuson–Stevens Fisheries Conservation and Management Reauthorization Act of 2006).

This study examined recall error in the FES by evaluating the impacts of bounded recall and the length of the reference period on reports of recreational saltwater fishing trips. We compared FES estimates of shore and private boat fishing effort to estimates derived from two experimental designs: one in which respondents were asked to report fishing trips for a single month (i.e., a 1-month reference period); and another that asked respondents to recall fishing trips for each of two separate months (i.e., reporting for the most recent month, bounded by reporting for the prior month). All design elements other than the reference period were identical between the FES and experimental treatments in an effort to minimize confounding effects. In comparing results from the experimental surveys, we explored possible mechanisms for any suspected recall biases.

## METHODS

*Experimental design.*— The FES is administered at the end of 2-month, mutually exclusive reference periods and asks respondents to recall the cumulative number of shore and private boat fishing trips that occurred during the reference period. From July to December 2015 (Table 1), two experimental questionnaires, which differed from the FES in the duration of the reference period, were administered in parallel to the FES in four states (Massachusetts, Maryland, Georgia, and Florida). One treatment (treatment 1) asked about fishing trips for two individual months (the most recent month and the prior month). The second treatment (treatment 2) asked about fishing trips for only the most recent month (see Appendix Figure A.1.1 for the differences between FES, treatment 1, and treatment 2 questionnaires). The experimental treatments were feasible modifications to the FES design that would provide greater temporal resolution and might potentially improve the accuracy of survey estimates.

With the exception of the manipulation of reference periods, the design of the FES and that of the experimental treatments were the same (Figure 1). The sample frame for each survey was the U.S. Postal Service's computerized delivery sequence file, consisting of all residential household addresses within each study state. The Massachusetts, Maryland, and Georgia samples were stratified into substate regions (groups of counties) defined by geographic proximity to the coast (coastal and noncoastal), while all counties in Florida were included in a single stratum due to the relatively high rate of fishing throughout the state. Within the geographic strata, we selected addresses using simple random sampling and matched them to the National Saltwater Angler Registry (MRIP 2018). This partitioned the sample into two additional strata: license matched (wherein the households contain one or more licensed anglers) and license unmatched (wherein no licensed anglers were identified in the household). This stratification provided additional information to optimize sampling; previous studies (e.g., Andrews et al. 2010, 2013; Brick et al. 2012) have demonstrated that residents of households that match to license databases respond to fishing surveys at a higher rate and are more likely to have fished during the reference wave than residents of unmatched households.

TABLE 1. Data collection schedule for the Fishing Effort Survey (FES; 2-month reference period), experimental treatment 1 (T1; both 1- and 2-month reference periods), and experimental treatment 2 (T2; 1-month reference period). Survey questionnaires were mailed out for the FES every 2 months (at the end of August, October, and December). Treatment 1 questionnaires were mailed out monthly from August to December. Treatment 2 questionnaires were sent out monthly from July to December.

| | Experimental month | | | | | |
|---|---|---|---|---|---|---|
| Variable or event | Jul | Aug | Sep | Oct | Nov | Dec |
| Treatment | T2 | FES, T1, T2 | T1, T2 | FES, T1, T2 | T1, T2 | FES, T1, T2 |
| First survey mailing | Jul 27, 2015 | Aug 25, 2015 | Sep 24, 2015 | Oct 26, 2015 | Nov 24, 2015 | Dec 28, 2015 |
| Reminder postcard | Aug 3, 2015 | Sep 1, 2015 | Oct 1, 2015 | Nov 2, 2015 | Dec 1, 2015 | Jan 4, 2016 |
| Reminder phone call | Aug 6, 2015 | Sep 3, 2015 | Oct 2, 2015 | Nov 4, 2015 | Dec 2, 2015 | Jan 4, 2016 |
| Second survey mailing | Aug 17, 2015 | Sep 15, 2015 | Oct 15, 2015 | Nov 16, 2015 | Dec 15, 2015 | Jan 18, 2016 |

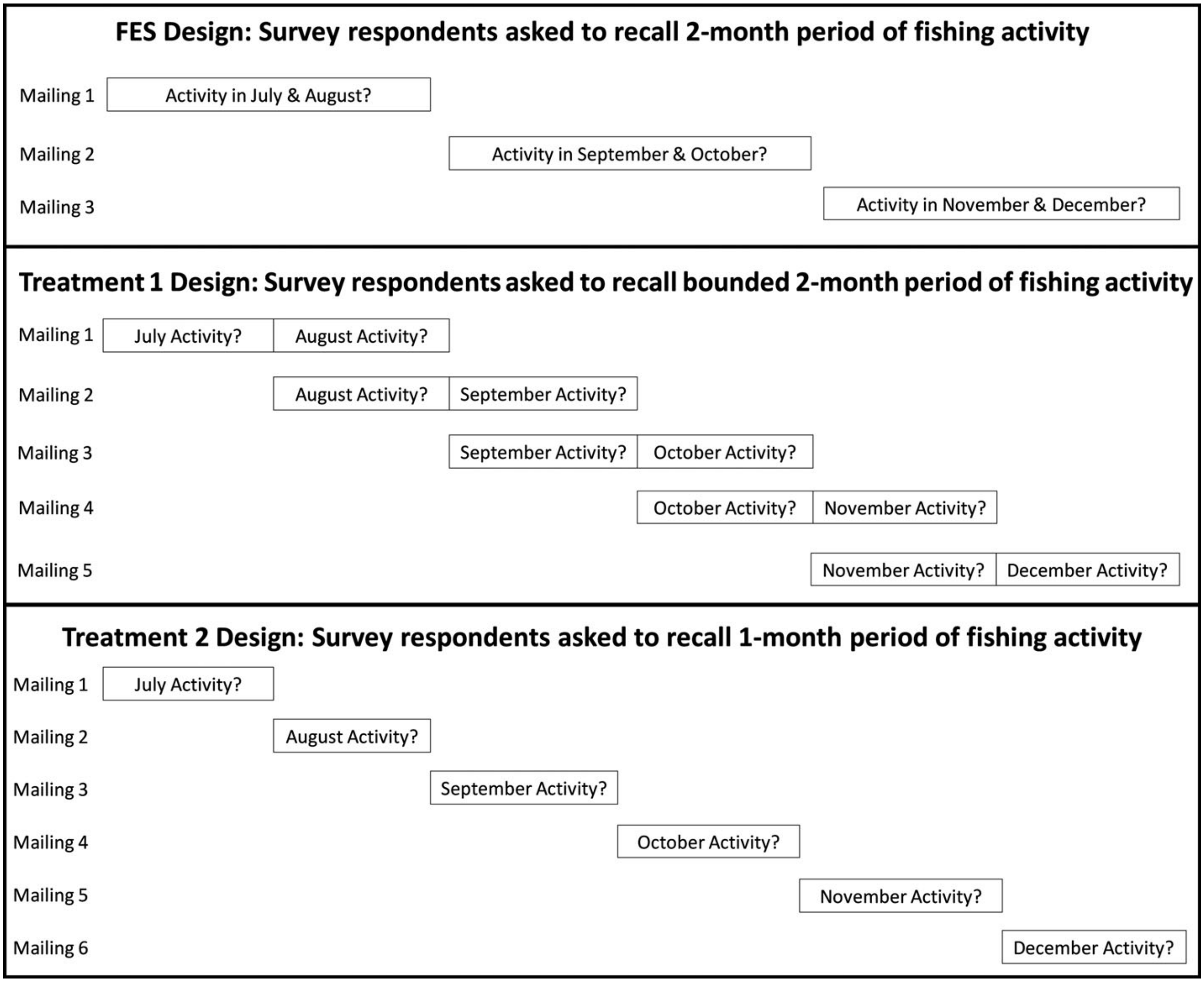



FIGURE 1. Schematic of the Fishing Effort Survey (FES) design compared to those of the experimental surveys: the FES was administered every 2 months and had a 2-month reference period (i.e., time frame for which survey respondents were asked to report events). Treatment 1 was administered monthly; respondents were given two reference periods and were asked to differentiate between fishing trips that occurred within the past month (1 month ago) and the month prior to that (2 months ago). Treatment 2 was administered monthly with a 1-month reference period.

The final sampling allocation was achieved by retaining all license-matched addresses in the sample and subsampling unmatched addresses at a rate of approximately 30%. The assignment to experimental treatments was completed after matching and subsampling; addresses within each stratum were randomly assigned to receive one of the two experimental versions of the survey. Sampling for the FES was conducted independently from the experimental treatments. In total, 39,539 questionnaires were mailed (Table 2), including treatment 1 (11,983 questionnaires), treatment 2 (12,017 questionnaires), and the FES (15,539 questionnaires). Table 2 presents sample sizes by state, and Appendix Table A.2.1 presents sample sizes by stratum.

*Data collection.*— Reported saltwater fishing trips were collected from occupants of each sampled address (up to a maximum of five household members) through a self-administered questionnaire. The data collection period began 1 week prior to the end of the reference month with an initial survey mailing that included a cover letter stating the purpose of the survey, a survey questionnaire, a postage-paid business reply envelope, and a prepaid US$2 cash incentive. One week after the initial mailing, households received an automated voice telephone reminder

TABLE 2. Sample sizes and responses by state for the 6-month experimental period (FES = Fishing Effort Survey). For a more detailed breakdown of sample sizes and responses for individual strata (i.e., by month, state, geographic stratum [coastal/noncoastal], and license status [matched/unmatched]), see Table A.2.1.

| | Treatment 1 | | Treatment 2 | | FES | |
|---|---|---|---|---|---|---|
| State | Initial sample size | Responses | Initial sample size | Responses | Initial sample size | Responses |
| Florida | 2,998 | 961 | 3,002 | 999 | 1,590 | 527 |
| Georgia | 2,995 | 988 | 3,005 | 974 | 4,244 | 1,402 |
| Maryland | 2,994 | 1,043 | 3,006 | 1,062 | 5,564 | 1,968 |
| Massachusetts | 2,996 | 1,142 | 3,004 | 1,062 | 4,141 | 1,554 |

message and a thank you/reminder postcard. Three weeks after the initial mailing, households received a second questionnaire, a nonresponse conversion letter designed to persuade nonresponding households to participate in the survey (Olson et al. 2011), and another postage-paid business reply envelope (see Table 1 for the data collection schedule for the experiment). Data were collected for approximately 13 weeks after the initial survey mailing for each reference month.

*Fishing effort estimation.—* Initial comparisons were of total shore and private boat fishing effort across the four experimental states for the entire 6-month experimental period. However, given the large influence of Florida, which accounted for approximately 75% of total effort for the four experimental states, we decided to consider Florida separately from the three other states. We considered shore and private boat fishing separately because the activities can be very different in terms of cost and time commitments—two factors that are likely to impact memory. Both treatment 1 and treatment 2 estimates were based upon the month immediately preceding survey administration; for treatment 1, this coincided with the most recent month of the 2-month reference period.

Initially, we compared FES trip estimates to experimental estimates to evaluate the impact of the different reference periods on survey estimates. Specifically, we hoped to determine whether estimates derived from a longer reference period were susceptible to recall decay. Next, we compared the experimental estimates to each other. We expected estimates from the two treatments to be similar since both were based upon reported fishing activity during the most recent month. Differences between treatments would presumably reflect the impact of the bounded recall design—asking about a behavior for multiple periods—on reporting. In addition to comparing the estimated number of trips across experimental treatments, we also compared fishing prevalence (percentage of households that reported fishing) and the mean number of trips reported per fishing household. Differences in these measures could help identify a mechanism for recall errors (Table 3).

Fishing prevalence and mean trips per household were calculated for treatments 1 and 2 by using established weighted mean estimators (SAS Institute 2016). Estimates of total fishing effort ($\widehat{T}_r$) for the FES, treatment 1, and treatment 2 were generated using the Horvitz–Thompson total estimator, a standard method for estimating the total of a stratified sample (Horvitz and Thompson 1952),

$$\widehat{T}_r = \sum_{h=1}^{H} \sum_{i=1}^{n_h} w_{hi} t_{hi},$$

where $w_{hi}$ is the weight of address $i$ in stratum $h$; and $t_{hi}$ is the reported number of recreational fishing trips for address $i$ in stratum $h$. The sample weights ($w_{hi}$) were calculated in a series of four steps that included (1) a base weight reflecting the sample inclusion probability; (2) an adjustment to account for unit nonresponse; (3) a post-stratification adjustment to account for incomplete coverage of the target population (e.g., Brick and Kalton 1996) using the most recent, reliable estimates of the number of residential households available from the American Community Survey (U.S. Census Bureau 2015) as population controls; and (4) use of an established procedure for trimming the estimated mean square error (see Potter 1990) to minimize the effects of extreme weights on the sampling variance.

The variance of the fishing effort estimates was calculated using Taylor series linearization (Dienes 1957; SAS Institute 2016). The Taylor series obtains a linear approximation of a nonlinear function, and the variance estimate of the nonlinear function is then estimated by the variance of the Taylor series approximation of that function (Woodruff 1971; Fuller 1975). The method calculates the estimated variance as

$$\widehat{V}(\widehat{T}_r) = \sum_{h=1}^{H} \left[ \frac{n_h}{n_h - 1} \left( \sum_{i=1}^{n_h} w_{hi}^* t_{hi} - \frac{1}{n_h} \sum_{i=1}^{n_h} w_{hi}^* t_{hi} \right)^2 \right].$$

## RESULTS

Of the over 10,000 questionnaires mailed for each of the experimental treatments, between 647 and 665 were undeliverable, and between 3,385 and 3,440 were completed and

TABLE 3. Statistical comparisons made between the survey estimates (FES = Fishing Effort Survey; T1, T2 = experimental treatments 1 and 2), along with the purpose of each, the expected outcomes, and potential mechanisms behind the expected outcomes.

| Comparison | Primary purpose of comparison | Expected outcome | Potential mechanisms |
|---|---|---|---|
| FES total effort to T1 total effort | Identify recall decay in the longer reference period (FES) | FES estimates lower than T1 estimates | Recall decay in the FES. |
| FES total effort to T2 total effort | Identify recall decay in the longer reference period (FES) | FES estimates lower than T2 estimates | Recall decay in the FES or telescoping in T2. |
| T1 total effort to T2 total effort | Examine the impact of a bounded recall design (T1) on estimates | Comparable estimates with no systematic differences | No difference in recall because the reference periods are the same (if T2 estimates are instead higher than T1 estimates, it would suggest telescoping in T2). |
| T1 fishing prevalence to T2 fishing prevalence | Explore mechanisms of observed recall error | Comparable estimates with no systematic differences | No difference in recall because the reference periods are the same (differences between treatments suggest that recall error is likely due to nonfishing households erroneously reporting fishing activity, thus indicating telescoping, social desirability, or a combination of both factors). |
| T1 mean trips per household to T2 mean trips per household | Explore mechanisms of observed recall error | Comparable estimates with no systematic differences | No difference in recall because the reference periods are the same (differences between treatments suggest that recall error is likely due to fishing households over- or underestimating the number of trips they took, indicating that recall ability is impacted by the frequency/regularity of fishing activity). |

returned (see Table 2 for responses by state and Table A.2.1 for responses by stratum). Of the nearly 16,000 FES questionnaires that were mailed during the 6-month experimental period, 745 were undeliverable, and 5,657 were returned. Adjusted response rates across all surveys were very similar, ranging from 36.21% to 37.25%.

Differences in estimated fishing trips between the FES and the two experimental treatments were not statistically significant for either shore or private boat fishing. However, treatment 2 estimates were systematically higher than FES estimates for both fishing modes (Figure 2). In contrast, differences between FES and treatment 1 estimates were neither significant nor systematic (Figure 2).

Comparisons between the experimental treatments demonstrated that treatment 2 trip estimates were systematically higher than treatment 1 estimates for both fishing modes (Figure 2). Differences between treatments were significant ($P < 0.05$) for both shore and private boat fishing in Florida and for private boat fishing in the remaining states. Differences in trip estimates resulted from differences in fishing prevalence between the two treatments; a higher percentage of households reported fishing when the overall reference period was limited to a single month (Figure 3). Differences in fishing prevalence between treatments 1 and 2 were significant for both shore and private boat fishing in Florida as well as for private boat fishing in the other states. In contrast, differences between treatments in terms of the mean trips per household were relatively minor and not significant (Figure 4).

## DISCUSSION

The FES estimates of total fishing effort were not significantly different from experimental estimates derived from a 1-month recall period (either treatment 1 or treatment 2). However, FES estimates were systematically lower than experimental estimates when the recall period was limited to a single month (treatment 2). This could mean that FES respondents were forgetting or omitting trips from the longer (2-month) recall period, resulting in

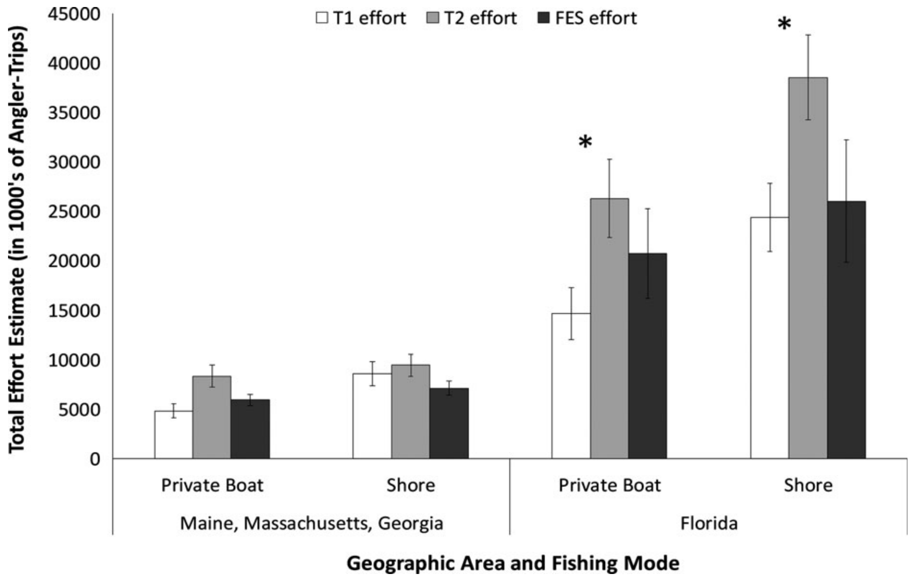



FIGURE 2. Comparison of fishing effort estimates (±SE; thousands of trips) from treatments 1 (T1) and 2 (T2) to each other and to the Fishing Effort Survey (FES) estimates by geographic area and by fishing mode. Estimates for each treatment were calculated for each reference period (T1 used 1-month estimates derived from the most recent month in the treatment's 2-month period; T2 used 1-month estimates; and the FES used 2-month estimates) and were summed across the 6-month experimental period. There were no significant differences in total fishing effort between the FES and either T1 or T2 ($P > 0.05$). Significant differences between T1 and T2 estimates are indicated by asterisks ($P < 0.05$).

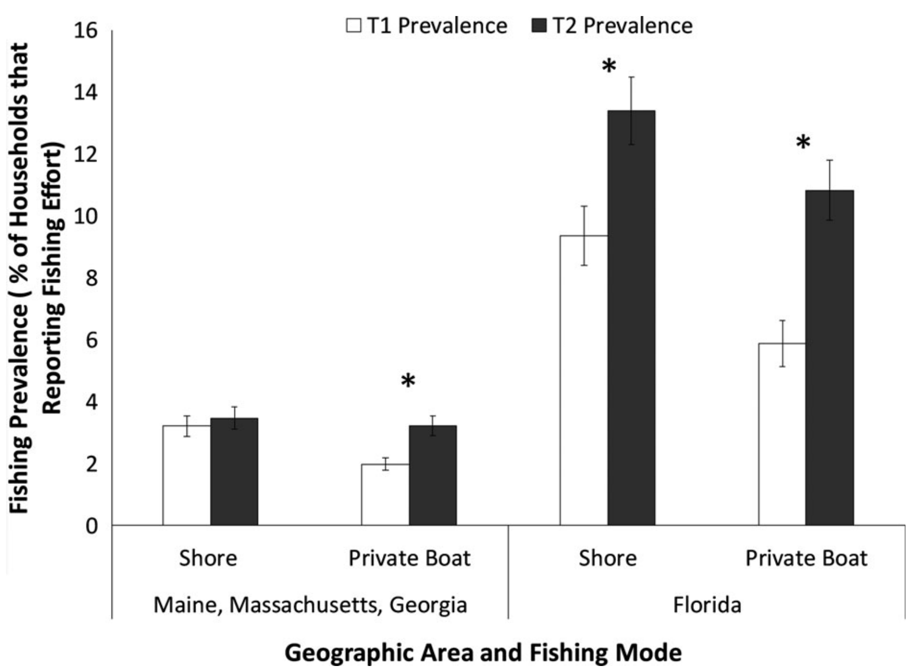



FIGURE 3. Comparison of fishing prevalence (±SE; percentage of households reporting fishing) in treatment 1 (T1; using the most recent of the 2 months within the treatment) and treatment 2 (T2; 1-month reference period) by geographic area and fishing mode. Significant differences between T1 and T2 metrics are indicated by asterisks ($P < 0.05$).

moderate underestimates of fishing effort. If this was the case, we would also expect FES estimates to be lower than estimates derived from the most recent month of a 2-month reference period (treatment 1). Differences between FES and treatment 1 estimates were neither significant nor systematic, suggesting that differences between FES and

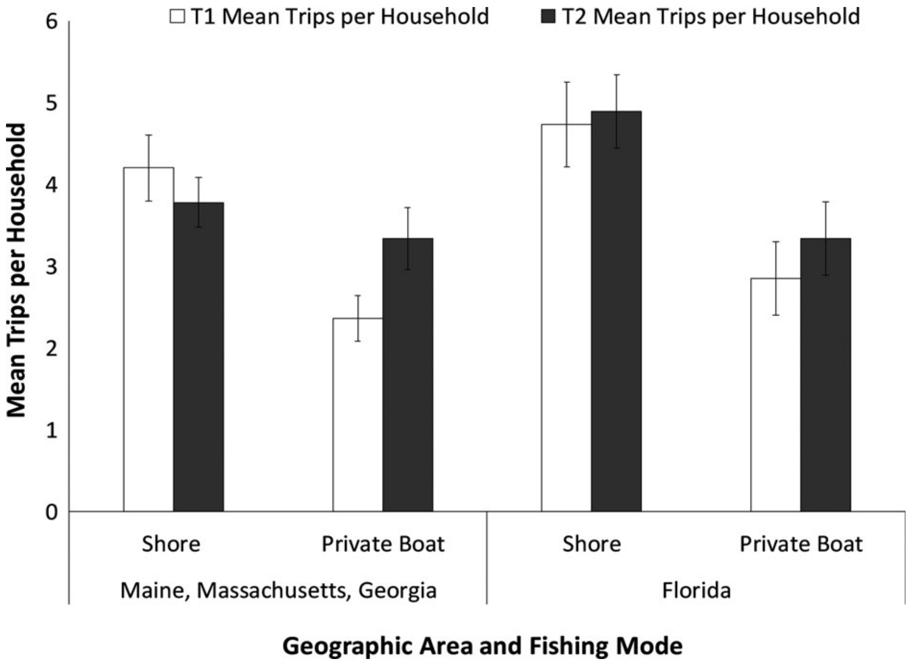



FIGURE 4. Comparison of mean fishing trips per household (±SE) in treatment 1 (T1; using the most recent of the 2 months within the treatment) and treatment 2 (T2; 1-month reference period) by geographic area and fishing mode. There were no significant differences in mean trips per household between T1 and T2 ($P > 0.05$).

treatment 2 estimates were not the result of omission error in the FES.

An alternative explanation for the differences between FES estimates and those based upon a single month (treatment 2) is that when asked to report for a single month, respondents telescoped trips from prior months into the reference period. This explanation is consistent with the observed differences between treatment 2 and treatment 1 estimates, both of which were based upon reported fishing trips during the most recent month and had the same recall period. The distinction between treatments 1 and 2 was that treatment 1 utilized a bounded design, asking first about fishing activity during the more distant month before asking about the more recent month.

Differences between trip estimates from treatments 1 and 2 were the result of differences in fishing prevalence rather than differences in the number of trips reported per household: more households reported fishing when the reference period was limited to a single month, but those households that did report fishing reported a similar number of trips, regardless of treatment. This result may reflect social desirability bias (Chu et al. 1989) or the desire by respondents to complete the requested task of reporting some level of fishing effort (Sudman and Bradburn 1974). In other words, respondents may think they are being helpful by providing a positive response to questions about fishing effort. Anglers who actually did fish are able to satisfy this desire without having to telescope trips into the reference period. The longer FES reference period may help to satisfy this desire and may partially mitigate the impacts of telescoping error by increasing the probability that a respondent actually did fish during the reference period.

Similarly, asking about fishing trips for two separate months, as in treatment 1, may minimize telescoping error for the most recent month by providing bounds against which responses are based. Neter and Waksberg (1964), who utilized a panel approach to improve recall and minimize telescoping error, initially described the potential benefits of bounded recall. In their design, the initial interview provided a recall bound for subsequent interviews. Sudman et al. (1984) modified the design to apply bounded interviewing in a single contact by asking about behaviors for multiple periods—first an earlier period and then a more recent period. Sudman et al. (1984) and others (Loftus et al. 1990) found that this approach reduced telescoping in the more recent reference period, resulting in lower, more accurate estimates. Our results suggest that bounded recall (as in treatment 1) minimizes telescoping for the most recent reference month by providing an additional opportunity for respondents to report a socially desirable behavior.

Based upon the results from this study, we cannot attribute differences in estimates between the FES and experimental treatments to recall error in the FES design. In fact, limiting the recall period to a single month appeared to increase recall error, resulting in overestimates of fishing effort. These results were consistent across geographic regions and fishing modes. If shorter, 1-month estimates are desired, however, our results suggest that a bounded 2-month design may be optimal for reducing recall error by using data from the second, most recent month of the reference period. These findings highlight the need for careful consideration of changes to survey designs, as subtle questionnaire differences can have substantial impacts on survey results. In weighing the trade-offs of survey design changes, consideration must also be given to precision, the subsequent sampling requirements needed to support different levels of resolution, and the impact of increased sampling on survey costs.

## ACKNOWLEDGMENTS

We thank Gallup for administering data collection for the FES and our experimental surveys as well as all of the participants that took the time to respond. We also thank Ryan Kitts-Jensen for useful feedback during the initial development of the project. We appreciate Geraldine Vander Haegen and two anonymous reviewers for providing helpful comments on the manuscript. There is no conflict of interest declared in this article.

## ORCID

Katherine J. Papacostas https://orcid.org/0000-0002-2937-4968

## REFERENCES

Abbott, W. F., and R. J. Monsen. 1979. On the measurement of corporate social responsibility: self-reported disclosures as a method of measuring corporate social involvement. Academy of Management Journal 22:501–515.

Andrews, R., J. M. Brick, and N. A. Mathiowetz. 2013. Continued development and testing of dual-frame surveys of fishing effort: testing a dual-frame, mixed-mode survey design. National Oceanic and Atmospheric Administration Fisheries, Office of Science and Technology, Silver Spring, Maryland.

Andrews, R., J. M. Brick, and N. A. Mathiowetz. 2014. Development and testing of recreational fishing effort surveys: testing a mail survey design. National Oceanic and Atmospheric Administration Fisheries, Office of Science and Technology, Silver Spring, Maryland.

Andrews, W. R., J. M. Brick, N. M. Mathiowetz, and L. Stokes. 2010. Pilot test of a dual-frame two-phase mail survey of anglers in North Carolina. National Oceanic and Atmospheric Administration Fisheries, Office of Science and Technology, Silver Spring, Maryland.

Bhandari, A., and T. Wagner. 2006. Self-reported utilization of health care services: improving measurement and accuracy. Medical Care Research and Review 63:217–235.

Blair, E. A., and S. Burton. 1987. Cognitive processes used by survey respondents to answer behavioral frequency questions. Journal of Consumer Research 14:280–288.

Brick, J., and G. Kalton. 1996. Handling missing data in survey research. Statistical Methods in Medical Research 5:215–238.

Brick, J. M., W. R. Andrews, and N. A. Mathiowetz. 2012. A comparison of recreational fishing effort survey designs. National Oceanic and Atmospheric Administration Fisheries, Office of Science and Technology, Silver Spring, Maryland.

Chu, A., D. Eisenhower, M. Hay, D. Morganstein, J. Neter, and J. Waksberg. 1992. Measuring the recall error in self-reported fishing and hunting activities. Journal of Official Statistics 8:19–39.

Chu, A., D. Eisenhower, D. Morganstein, and J. Waksberg. 1989. Investigation of possible recall/reference period bias in national surveys of fishing, hunting and wildlife-associated recreation. Report by Westat to the U.S. Fish and Wildlife Service, Contract 14-16-009-87-008, Washington, D.C.

Clarke, P. M., D. G. Fiebig, and U.-G. Gerdtham. 2008. Optimal recall length in survey design. Journal of Health Economics 27:1275–1284.

Connelly, N. A., and T. L. Brown. 1995. Use of angler diaries to examine biases associated with 12-month recall on mail questionnaires. Transactions of the American Fisheries Society 124:413–422.

Connelly, N. A., and T. L. Brown. 2011. Effect of recall period on annual freshwater fishing effort estimates in New York. Fisheries Management and Ecology 18:83–87.

Connelly, N. A., T. L. Brown, and B. A. Knuth. 2000. Assessing the relative importance of recall bias and nonresponse bias and adjusting for those biases in statewide angler surveys. Human Dimensions of Wildlife 5(4):19–29.

Dienes, P. 1957. The Taylor series: an introduction to the theory of functions of a complex variable. Dover Publications, New York.

Ditton, R. B., and K. M. Hunt. 2008. Combining creel intercept and mail survey methods to understand the human dimensions of local freshwater fisheries. Fisheries Management and Ecology 8:295–301.

Eisenhower, D., N. A. Mathiowetz, and D. Morganstein. 2011. Recall error: sources and bias reduction techniques. Pages 125–144 *in* P. P. Biemer, R. M. Groves, L. E. Lyberg, N. A. Mathiowetz, and S. Sudman, editors. Measurement error in surveys. Wiley, Hoboken, New Jersey.

Fuller, W. A. 1975. Regression analysis for sample survey. Sankhya, Series C 37:117–132.

Gaskell, G. D., D. B. Wright, and C. A. O'Muircheartaigh. 2000. Telescoping of landmark events—implications for survey research. Public Opinion Quarterly 64:77–89.

Gems, B., D. Ghosh, and R. Hitlin. 1982. A recall experiment: impact of time on recall of recreational fishing trips. Pages 372–375 *in* Proceedings of the survey research methods section. American Statistical Association, Alexandria, Virginia.

Hicks, R. L., A. B. Gautam, D. Van Voorhees, M. Osborn, and B. Gentner. 1999. An introduction to the NMFS Marine Recreational Fisheries Statistics Survey with an emphasis on economic valuation. Marine Resource Economics 14:375–385.

Horvitz, D. G., and D. J. Thompson. 1952. A generalization of sampling without replacement from a finite universe. Journal of the American Statistical Association 47:663–685.

Loftus, E. F., M. R. Klinger, K. D. Smith, and J. Fiedler. 1990. A tale of two questions: benefits of asking more than one question. Public Opinion Quarterly 54:330–345.

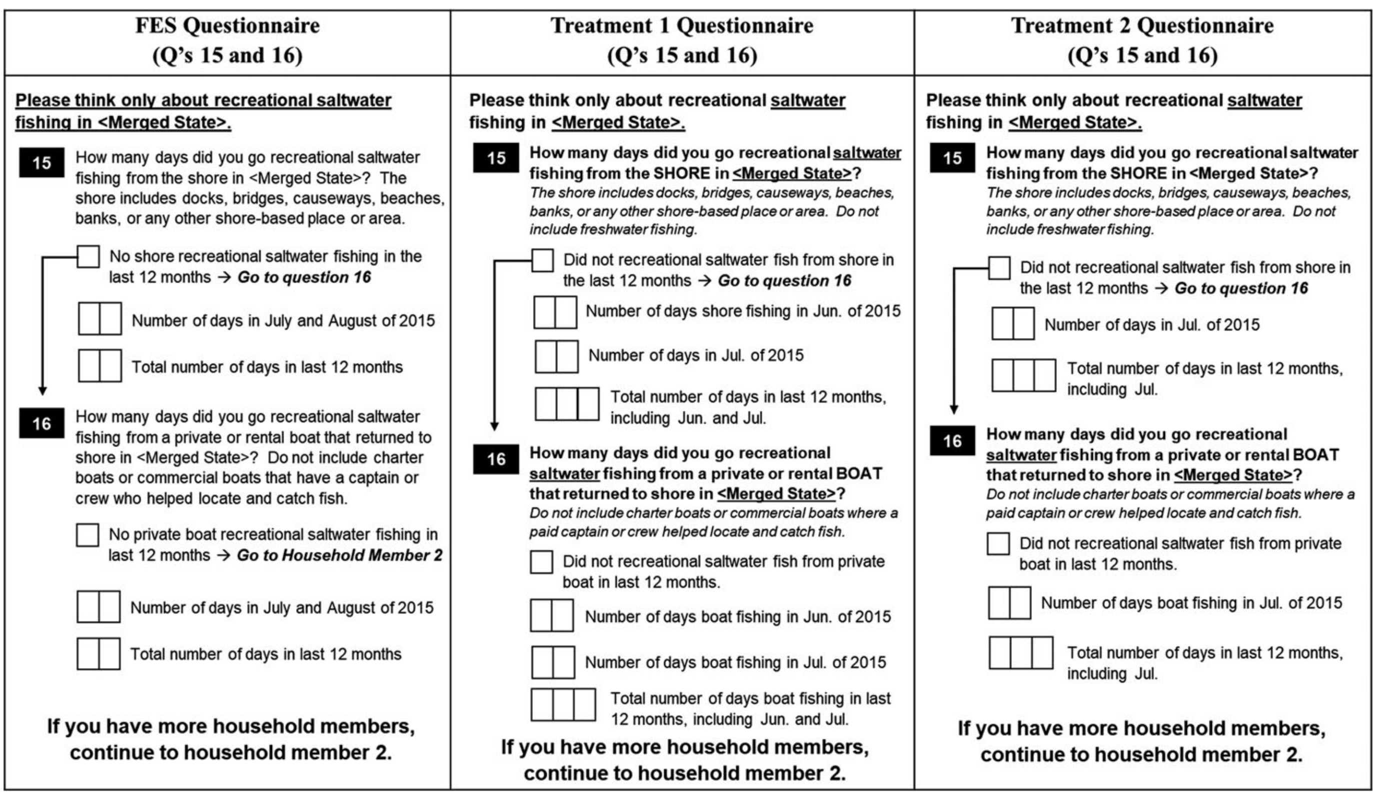
Loftus, E. F., and W. Marburger. 1983. Since the eruption of Mt. St. Helens, has anyone beaten you up? Improving the accuracy of retrospective reports with landmark events. Memory & Cognition 11:114–120.

Magnuson–Stevens Fisheries Conservation and Management Reauthorization Act. 2006. National Standard 2–Scientific Information. U.S Code, volume 50, section 600.315.

MRIP (Marine Recreational Information Program). 2018. National saltwater angler registry. National Oceanic and Atmospheric

Administration. Available: https://www.countmyfish.noaa.gov/register/. (November 2018).
National Academies of Sciences, Engineering, and Medicine. 2017. Review of the marine recreational information program. National Academies Press, Washington, D.C.
National Research Council. 2006. Review of recreational fisheries survey methods. National Academies Press, Washington, D.C.
Neter, J., and J. Waksberg. 1964. A study of response errors in expenditures data from household interviews. Journal of the American Statistical Association 59:18–55.
Olson, K., J. M. Lepkowski, D. H. Garabrant, and N. Lincoln Olson. 2011. An experimental examination of the content of persuasion letters on nonresponse rates and survey estimates in a nonresponse follow-up study. Survey Research Methods 5:21–26.
Potter, F. J. 1990. A study of procedures to identify and trim extreme sampling weights. Research Triangle Institute, Research Triangle Park, North Carolina.
Rocklin, D., H. Levrel, M. Drogou, J. Herfaut, and G. Veron. 2014. Combining telephone surveys and fishing catches self-report: the French Sea Bass Recreational Fishery Assessment. PLoS ONE [online serial] 9(1):e87271.
Sampson, D. B. 2011. The accuracy of self-reported fisheries data: Oregon trawl logbook fishing locations and retained catches. Fisheries Research 112:59–76.
SAS Institute. 2016. SAS/STAT 14.2 user's guide. SAS Institute, Cary, North Carolina.
Sudman, S., and N. M. Bradburn. 1973. Effects of time and memory factors on response in surveys. Journal of the American Statistical Association 68:805–815.
Sudman, S., and N. M. Bradburn. 1974. Response effects in surveys: a review and synthesis. National Opinion Research Center, Monographs in Social Research, Chicago.
Sudman, S., A. Finn, and L. Lannom. 1984. The use of bounded recall procedures in single interviews. Public Opinion Quarterly 48:520–524.
Tarrant, M. A., M. J. Manfredo, P. B. Bayley, and R. Hess. 1993. Effects of recall bias and nonresponse bias on self-report estimates of angling participation. North American Journal of Fisheries Management 13:217–222.
U.S. Census Bureau. 2015. American community survey 2014 data release. Available: https://www.census.gov/programs-surveys/acs/news/data-releases/2014/release.html#par_textimage_11. (November 2018).
Woodruff, R. S. 1971. A simple method for approximating the variance of a complicated estimate. Journal of the American Statistical Association 66:411–414.
Wright, E. R., and B. A. Pescosolido. 2002. "Sorry, I forgot": the role of recall error in longitudinal personal network studies. Pages 113–129 *in* J. A. Levy and B. A. Pescosolido, editors. Social networks and health: advances in medical sociology. Emerald Group Publishing, Bingley, UK.

## Appendix 1: Difference in Questionnaires



**FES Questionnaire (Q's 15 and 16)**

Please think only about recreational saltwater fishing in <Merged State>.

15 How many days did you go recreational saltwater fishing from the shore in <Merged State>? The shore includes docks, bridges, causeways, beaches, banks, or any other shore-based place or area.

No shore recreational saltwater fishing in the last 12 months → *Go to question 16*

Number of days in July and August of 2015

Total number of days in last 12 months

16 How many days did you go recreational saltwater fishing from a private or rental boat that returned to shore in <Merged State>? Do not include charter boats or commercial boats that have a captain or crew who helped locate and catch fish.

No private boat recreational saltwater fishing in last 12 months → *Go to Household Member 2*

Number of days in July and August of 2015

Total number of days in last 12 months

**If you have more household members, continue to household member 2.**

**Treatment 1 Questionnaire (Q's 15 and 16)**

Please think only about recreational saltwater fishing in <Merged State>.

15 **How many days did you go recreational saltwater fishing from the SHORE in <Merged State>?** *The shore includes docks, bridges, causeways, beaches, banks, or any other shore-based place or area. Do not include freshwater fishing.*

Did not recreational saltwater fish from shore in the last 12 months → *Go to question 16*

Number of days shore fishing in Jun. of 2015

Number of days in Jul. of 2015

Total number of days in last 12 months, including Jun. and Jul.

16 **How many days did you go recreational saltwater fishing from a private or rental BOAT that returned to shore in <Merged State>?** *Do not include charter boats or commercial boats where a paid captain or crew helped locate and catch fish.*

Did not recreational saltwater fish from private boat in last 12 months.

Number of days boat fishing in Jun. of 2015

Number of days boat fishing in Jul. of 2015

Total number of days boat fishing in last 12 months, including Jun. and Jul.

**If you have more household members, continue to household member 2.**

**Treatment 2 Questionnaire (Q's 15 and 16)**

Please think only about recreational saltwater fishing in <Merged State>.

15 **How many days did you go recreational saltwater fishing from the SHORE in <Merged State>?** *The shore includes docks, bridges, causeways, beaches, banks, or any other shore-based place or area. Do not include freshwater fishing.*

Did not recreational saltwater fish from shore in the last 12 months → *Go to question 16*

Number of days in Jul. of 2015

Total number of days in last 12 months, including Jul.

16 **How many days did you go recreational saltwater fishing from a private or rental BOAT that returned to shore in <Merged State>?** *Do not include charter boats or commercial boats where a paid captain or crew helped locate and catch fish.*

Did not recreational saltwater fish from private boat in last 12 months.

Number of days boat fishing in Jul. of 2015

Total number of days in last 12 months, including Jul.

**If you have more household members, continue to household member 2.**

FIGURE A.1.1. Difference among the treatment 1, treatment 2, and Fishing Effort Survey (FES) questionnaires. The questionnaires consisted of 16 questions for up to five people living in the household. The surveys differed only in questions 15 and 16, which were about recalling shore and private boat fishing activity. Questions 15 and 16 for each of the three surveys used in this study are presented.

## Appendix 2: Sample Sizes and Response Rates

TABLE A.2.1. Sample sizes and response rates per stratum and the estimated total number of households in each stratum for treatment 1 (T1), treatment 2 (T2), and the Fishing Effort Survey (FES).

| Survey | Month | State | Geographic stratum (coastal or noncoastal) | License status (matched or unmatched) | Sample size | Complete surveys | Estimated number of households in the stratum |
|---|---|---|---|---|---|---|---|
| T1 | Jul | FL | Coastal | Matched | 108 | 41 | 699,510 |
| | | FL | Coastal | Unmatched | 392 | 116 | 6,512,564 |
| | | GA | Coastal | Matched | 157 | 68 | 45,540 |
| | | GA | Coastal | Unmatched | 160 | 45 | 261,939 |
| | | GA | Noncoastal | Matched | 48 | 19 | 147,656 |
| | | GA | Noncoastal | Unmatched | 134 | 43 | 3,094,723 |
| | | MD | Coastal | Matched | 34 | 14 | 70,723 |
| | | MD | Coastal | Unmatched | 329 | 98 | 1,848,157 |
| | | MD | Noncoastal | Matched | 33 | 10 | 8,173 |
| | | MD | Noncoastal | Unmatched | 103 | 43 | 235,552 |
| | | MA | Coastal | Matched | 115 | 54 | 44,695 |
| | | MA | Coastal | Unmatched | 331 | 118 | 1,870,372 |
| | | MA | Noncoastal | Matched | 20 | 13 | 16,355 |
| | | MA | Noncoastal | Unmatched | 33 | 13 | 605,543 |
| T2 | Jul | FL | Coastal | Matched | 108 | 42 | 699,510 |
| | | FL | Coastal | Unmatched | 392 | 135 | 6,512,564 |
| | | GA | Coastal | Matched | 158 | 61 | 43,604 |
| | | GA | Coastal | Unmatched | 161 | 44 | 261,939 |
| | | GA | Noncoastal | Matched | 48 | 15 | 147,656 |
| | | GA | Noncoastal | Unmatched | 134 | 45 | 3,094,723 |
| | | MD | Coastal | Matched | 34 | 11 | 70,723 |
| | | MD | Coastal | Unmatched | 329 | 121 | 1,848,157 |
| | | MD | Noncoastal | Matched | 34 | 27 | 8,173 |
| | | MD | Noncoastal | Unmatched | 104 | 41 | 235,552 |
| | | MA | Coastal | Matched | 115 | 50 | 44,695 |
| | | MA | Coastal | Unmatched | 332 | 109 | 1,870,372 |
| | | MA | Noncoastal | Matched | 21 | 3 | 16,355 |
| | | MA | Noncoastal | Unmatched | 33 | 12 | 605,543 |
| T1 | Aug | FL | Coastal | Matched | 90 | 35 | 603,521 |
| | | FL | Coastal | Unmatched | 410 | 135 | 6,608,553 |
| | | GA | Coastal | Matched | 157 | 62 | 37,507 |
| | | GA | Coastal | Unmatched | 160 | 56 | 268,036 |
| | | GA | Noncoastal | Matched | 48 | 22 | 137,828 |
| | | GA | Noncoastal | Unmatched | 134 | 38 | 3,104,551 |
| | | MD | Coastal | Matched | 26 | 9 | 91,796 |
| | | MD | Coastal | Unmatched | 355 | 136 | 1,827,084 |
| | | MD | Noncoastal | Matched | 8 | 5 | 6,464 |
| | | MD | Noncoastal | Unmatched | 110 | 46 | 237,261 |
| | | MA | Coastal | Matched | 94 | 48 | 76,538 |
| | | MA | Coastal | Unmatched | 358 | 126 | 1,838,529 |
| | | MA | Noncoastal | Matched | 10 | 6 | 13,417 |
| | | MA | Noncoastal | Unmatched | 38 | 15 | 608,481 |

TABLE A.1. Continued.

| Survey | Month | State | Geographic stratum (coastal or noncoastal) | License status (matched or unmatched) | Sample size | Complete surveys | Estimated number of households in the stratum |
|---|---|---|---|---|---|---|---|
| T2 | Aug | FL | Coastal | Matched | 90 | 36 | 603,521 |
| | | FL | Coastal | Unmatched | 410 | 129 | 6,608,553 |
| | | GA | Coastal | Matched | 158 | 65 | 37,507 |
| | | GA | Coastal | Unmatched | 161 | 54 | 268,036 |
| | | GA | Noncoastal | Matched | 48 | 18 | 137,828 |
| | | GA | Noncoastal | Unmatched | 134 | 43 | 3,104,551 |
| | | MD | Coastal | Matched | 26 | 9 | 91,796 |
| | | MD | Coastal | Unmatched | 356 | 110 | 1,827,084 |
| | | MD | Noncoastal | Matched | 9 | 7 | 6,464 |
| | | MD | Noncoastal | Unmatched | 110 | 49 | 237,261 |
| | | MA | Coastal | Matched | 94 | 44 | 76,538 |
| | | MA | Coastal | Unmatched | 358 | 121 | 1,838,529 |
| | | MA | Noncoastal | Matched | 10 | 4 | 13,417 |
| | | MA | Noncoastal | Unmatched | 38 | 14 | 608,481 |
| FES | Jul/Aug | FL | Coastal | Matched | 74 | 24 | 647,686 |
| | | FL | Coastal | Unmatched | 309 | 96 | 6,564,388 |
| | | GA | Coastal | Matched | 359 | 155 | 47,275 |
| | | GA | Coastal | Unmatched | 366 | 110 | 268,962 |
| | | GA | Noncoastal | Matched | 109 | 43 | 141,962 |
| | | GA | Noncoastal | Unmatched | 305 | 96 | 3,089,723 |
| | | MD | Coastal | Matched | 60 | 22 | 86,113 |
| | | MD | Coastal | Unmatched | 879 | 326 | 1,832,767 |
| | | MD | Noncoastal | Matched | 20 | 8 | 7,593 |
| | | MD | Noncoastal | Unmatched | 272 | 100 | 236,132 |
| | | MA | Coastal | Matched | 158 | 86 | 67,843 |
| | | MA | Coastal | Unmatched | 699 | 264 | 1,847,224 |
| | | MA | Noncoastal | Matched | 25 | 15 | 16,754 |
| | | MA | Noncoastal | Unmatched | 66 | 20 | 605,144 |
| T1 | Sep | FL | Coastal | Matched | 101 | 46 | 680,637 |
| | | FL | Coastal | Unmatched | 398 | 123 | 6,531,437 |
| | | GA | Coastal | Matched | 157 | 66 | 39,333 |
| | | GA | Coastal | Unmatched | 160 | 49 | 266,210 |
| | | GA | Noncoastal | Matched | 48 | 22 | 122,817 |
| | | GA | Noncoastal | Unmatched | 134 | 41 | 3,119,562 |
| | | MD | Coastal | Matched | 29 | 9 | 102,387 |
| | | MD | Coastal | Unmatched | 352 | 118 | 1,816,493 |
| | | MD | Noncoastal | Matched | 10 | 5 | 7,593 |
| | | MD | Noncoastal | Unmatched | 108 | 40 | 236,132 |
| | | MA | Coastal | Matched | 91 | 40 | 74,142 |
| | | MA | Coastal | Unmatched | 361 | 125 | 1,840,925 |
| | | MA | Noncoastal | Matched | 15 | 8 | 20,797 |
| | | MA | Noncoastal | Unmatched | 32 | 9 | 601,101 |

TABLE A.1. Continued.

| Survey | Month | State | Geographic stratum (coastal or noncoastal) | License status (matched or unmatched) | Sample size | Complete surveys | Estimated number of households in the stratum |
|---|---|---|---|---|---|---|---|
| T2 | Sep | FL | Coastal | Matched | 102 | 31 | 680,637 |
| | | FL | Coastal | Unmatched | 399 | 129 | 6,531,437 |
| | | GA | Coastal | Matched | 158 | 64 | 39,333 |
| | | GA | Coastal | Unmatched | 161 | 52 | 266,210 |
| | | GA | Noncoastal | Matched | 48 | 14 | 122,817 |
| | | GA | Noncoastal | Unmatched | 134 | 39 | 3,119,562 |
| | | MD | Coastal | Matched | 29 | 11 | 102,387 |
| | | MD | Coastal | Unmatched | 353 | 117 | 1,816,493 |
| | | MD | Noncoastal | Matched | 10 | 2 | 7,593 |
| | | MD | Noncoastal | Unmatched | 109 | 38 | 236,132 |
| | | MA | Coastal | Matched | 91 | 32 | 74,142 |
| | | MA | Coastal | Unmatched | 361 | 112 | 1,840,925 |
| | | MA | Noncoastal | Matched | 16 | 9 | 20,797 |
| | | MA | Noncoastal | Unmatched | 33 | 12 | 601,101 |
| T1 | Oct | FL | Coastal | Matched | 96 | 34 | 648,276 |
| | | FL | Coastal | Unmatched | 404 | 109 | 6,563,798 |
| | | GA | Coastal | Matched | 140 | 40 | 38,814 |
| | | GA | Coastal | Unmatched | 177 | 65 | 266,729 |
| | | GA | Noncoastal | Matched | 64 | 26 | 116,218 |
| | | GA | Noncoastal | Unmatched | 117 | 25 | 3,126,161 |
| | | MD | Coastal | Matched | 21 | 9 | 90,872 |
| | | MD | Coastal | Unmatched | 360 | 108 | 1,828,008 |
| | | MD | Noncoastal | Matched | 10 | 7 | 8,928 |
| | | MD | Noncoastal | Unmatched | 108 | 39 | 234,797 |
| | | MA | Coastal | Matched | 76 | 38 | 74,315 |
| | | MA | Coastal | Unmatched | 376 | 133 | 1,840,752 |
| | | MA | Noncoastal | Matched | 11 | 2 | 19,729 |
| | | MA | Noncoastal | Unmatched | 36 | 14 | 602,169 |
| T2 | Oct | FL | Coastal | Matched | 96 | 33 | 648,276 |
| | | FL | Coastal | Unmatched | 404 | 134 | 6,563,798 |
| | | GA | Coastal | Matched | 141 | 49 | 38,814 |
| | | GA | Coastal | Unmatched | 178 | 46 | 266,729 |
| | | GA | Noncoastal | Matched | 65 | 20 | 116,218 |
| | | GA | Noncoastal | Unmatched | 118 | 30 | 3,126,161 |
| | | MD | Coastal | Matched | 22 | 10 | 90,872 |
| | | MD | Coastal | Unmatched | 360 | 118 | 1,828,008 |
| | | MD | Noncoastal | Matched | 10 | 4 | 8,928 |
| | | MD | Noncoastal | Unmatched | 109 | 40 | 234,797 |
| | | MA | Coastal | Matched | 76 | 37 | 74,315 |
| | | MA | Coastal | Unmatched | 376 | 123 | 1,840,752 |
| | | MA | Noncoastal | Matched | 12 | 5 | 19,729 |
| | | MA | Noncoastal | Unmatched | 37 | 13 | 602,169 |

TABLE A.1. Continued.

| Survey | Month | State | Geographic stratum (coastal or noncoastal) | License status (matched or unmatched) | Sample size | Complete surveys | Estimated number of households in the stratum |
|---|---|---|---|---|---|---|---|
| FES | Sep/Oct | FL | Coastal | Matched | 92 | 34 | 725,942 |
| | | FL | Coastal | Unmatched | 336 | 112 | 6,486,132 |
| | | GA | Coastal | Matched | 214 | 85 | 42,708 |
| | | GA | Coastal | Unmatched | 226 | 71 | 262,835 |
| | | GA | Noncoastal | Matched | 94 | 29 | 121,331 |
| | | GA | Noncoastal | Unmatched | 158 | 39 | 3,121,048 |
| | | MD | Coastal | Matched | 67 | 31 | 108,769 |
| | | MD | Coastal | Unmatched | 926 | 289 | 1,810,111 |
| | | MD | Noncoastal | Matched | 33 | 20 | 14,731 |
| | | MD | Noncoastal | Unmatched | 276 | 103 | 228,994 |
| | | MA | Coastal | Matched | 138 | 58 | 67,227 |
| | | MA | Coastal | Unmatched | 772 | 280 | 1,846,442 |
| | | MA | Noncoastal | Matched | 27 | 7 | 23,117 |
| | | MA | Noncoastal | Unmatched | 69 | 19 | 600,180 |
| T1 | Nov | FL | Coastal | Matched | 89 | 29 | 604,383 |
| | | FL | Coastal | Unmatched | 410 | 116 | 6,607,691 |
| | | GA | Coastal | Matched | 146 | 57 | 31,139 |
| | | GA | Coastal | Unmatched | 172 | 42 | 204,749 |
| | | GA | Noncoastal | Matched | 63 | 18 | 115,953 |
| | | GA | Noncoastal | Unmatched | 119 | 32 | 3,196,081 |
| | | MD | Coastal | Matched | 25 | 13 | 105,665 |
| | | MD | Coastal | Unmatched | 356 | 110 | 1,813,215 |
| | | MD | Noncoastal | Matched | 12 | 6 | 10,713 |
| | | MD | Noncoastal | Unmatched | 106 | 43 | 233,012 |
| | | MA | Coastal | Matched | 77 | 37 | 75,292 |
| | | MA | Coastal | Unmatched | 375 | 118 | 1,839,775 |
| | | MA | Noncoastal | Matched | 9 | 2 | 15,440 |
| | | MA | Noncoastal | Unmatched | 39 | 13 | 606,458 |
| T2 | Nov | FL | Coastal | Matched | 90 | 35 | 604,383 |
| | | FL | Coastal | Unmatched | 411 | 124 | 6,607,691 |
| | | GA | Coastal | Matched | 146 | 54 | 31,139 |
| | | GA | Coastal | Unmatched | 172 | 45 | 204,749 |
| | | GA | Noncoastal | Matched | 63 | 22 | 115,953 |
| | | GA | Noncoastal | Unmatched | 119 | 25 | 3,196,081 |
| | | MD | Coastal | Matched | 25 | 13 | 105,665 |
| | | MD | Coastal | Unmatched | 357 | 112 | 1,813,215 |
| | | MD | Noncoastal | Matched | 12 | 8 | 10,713 |
| | | MD | Noncoastal | Unmatched | 107 | 38 | 233,012 |
| | | MA | Coastal | Matched | 77 | 33 | 75,292 |
| | | MA | Coastal | Unmatched | 375 | 114 | 1,839,775 |
| | | MA | Noncoastal | Matched | 9 | 5 | 15,440 |
| | | MA | Noncoastal | Unmatched | 39 | 17 | 606,458 |

TABLE A.1. Continued.

| Survey | Month | State | Geographic stratum (coastal or noncoastal) | License status (matched or unmatched) | Sample size | Complete surveys | Estimated number of households in the stratum |
|---|---|---|---|---|---|---|---|
| T1 | Dec | FL | Coastal | Matched | 92 | 30 | 628,325 |
| | | FL | Coastal | Unmatched | 408 | 147 | 6,583,749 |
| | | GA | Coastal | Matched | 120 | 46 | 33,030 |
| | | GA | Coastal | Unmatched | 198 | 51 | 202,858 |
| | | GA | Noncoastal | Matched | 43 | 16 | 118,533 |
| | | GA | Noncoastal | Unmatched | 139 | 39 | 3,193,501 |
| | | MD | Coastal | Matched | 15 | 11 | 68,610 |
| | | MD | Coastal | Unmatched | 366 | 112 | 1,850,270 |
| | | MD | Noncoastal | Matched | 17 | 9 | 16,187 |
| | | MD | Noncoastal | Unmatched | 101 | 43 | 227,538 |
| | | MA | Coastal | Matched | 84 | 44 | 84,902 |
| | | MA | Coastal | Unmatched | 367 | 143 | 1,830,165 |
| | | MA | Noncoastal | Matched | 11 | 8 | 17,955 |
| | | MA | Noncoastal | Unmatched | 37 | 15 | 603,943 |
| T2 | Dec | FL | Coastal | Matched | 92 | 42 | 628,325 |
| | | FL | Coastal | Unmatched | 408 | 129 | 6,583,749 |
| | | GA | Coastal | Matched | 120 | 45 | 33,030 |
| | | GA | Coastal | Unmatched | 198 | 70 | 202,858 |
| | | GA | Noncoastal | Matched | 43 | 15 | 118,533 |
| | | GA | Noncoastal | Unmatched | 139 | 39 | 3,193,501 |
| | | MD | Coastal | Matched | 16 | 3 | 68,610 |
| | | MD | Coastal | Unmatched | 366 | 124 | 1,850,270 |
| | | MD | Noncoastal | Matched | 18 | 13 | 16,187 |
| | | MD | Noncoastal | Unmatched | 101 | 36 | 227,538 |
| | | MA | Coastal | Matched | 85 | 45 | 84,902 |
| | | MA | Coastal | Unmatched | 368 | 130 | 1,830,165 |
| | | MA | Noncoastal | Matched | 11 | 5 | 17,955 |
| | | MA | Noncoastal | Unmatched | 37 | 13 | 603,943 |
| FES | Nov/Dec | FL | Coastal | Matched | 157 | 75 | 694,039 |
| | | FL | Coastal | Unmatched | 622 | 186 | 6,518,035 |
| | | GA | Coastal | Matched | 564 | 215 | 32,190 |
| | | GA | Coastal | Unmatched | 970 | 264 | 203,263 |
| | | GA | Noncoastal | Matched | 235 | 94 | 132,273 |
| | | GA | Noncoastal | Unmatched | 644 | 201 | 3,180,195 |
| | | MD | Coastal | Matched | 116 | 58 | 80,165 |
| | | MD | Coastal | Unmatched | 2,196 | 735 | 1,840,010 |
| | | MD | Noncoastal | Matched | 36 | 22 | 12,231 |
| | | MD | Noncoastal | Unmatched | 683 | 254 | 230,199 |
| | | MA | Coastal | Matched | 395 | 198 | 76,860 |
| | | MA | Coastal | Unmatched | 1,941 | 712 | 1,838,207 |
| | | MA | Noncoastal | Matched | 25 | 12 | 17,995 |
| | | MA | Noncoastal | Unmatched | 221 | 81 | 603,903 |

# Section 2: Evaluating Measurement Error in the MRIP Fishing Effort Survey: The Effect of Question Sequence on Reporting of Fishing Activity

# Evaluating Measurement Error in the MRIP Fishing Effort Survey: The Effect of Question Sequence on Reporting of Fishing Activity

**Rob Andrews**

**September 1, 2022**

## Background

The MRIP Fishing Effort Survey (FES) is a self-administered, household mail survey that was designed to estimate marine recreational shore and private boat fishing activity (Papacostas and Foster 2018). The survey is administered for discrete, two-month reference waves in the coastal states along the Atlantic coast and Gulf of Mexico, as well as Hawaii. The FES was developed over several years through a series of pilot studies that evaluated different sample frames, data collection modes, survey instruments and sample designs (Andrews et al. 2014). The goal of this testing was to develop a cost-effective survey design that minimized the potential for non-sampling errors. The final FES design has been endorsed by peer review and the National Academies of Science Engineering and Medicine (Andrews 2014, NASEM 2017).

### *Measurement Error*

A primary objective of FES testing was to design a straightforward questionnaire that reduced the potential for bias resulting from measurement error, specifically recall error. Recall error occurs when respondents provide inaccurate responses to survey questions (Neter and Waksberg 1964). Recall errors are generally classified as omission errors - when respondents forget events (also referred to as recall decay) - or telescoping errors - when respondents misplace events in time (Sudman and Bradburn 1973). Generally, omission error and telescoping error result in underestimates and overestimates of survey measures, respectively. Survey methodologists employ several techniques to minimize recall error, including adjusting the recall period, providing memory cues, associating the recall period with landmark events, and bounding recall against a prior survey request or question within the same survey (Loftus and Marburger 1983, Gaskett et al. 2000, Chu et al. 1989, Neter and Waksberg 1964, Sudman et al. 1984).

A range of studies were conducted to investigate and minimize possible measurement errors in the FES design. These studies started with cognitive interviews and progressed to split-ballot experiments to evaluate the effects of questionnaire design features on reporting of fishing activity. The split-ballot studies described below were designed specifically to evaluate the potential for both omission and telescoping errors.

In initially developing the FES questionnaire, care was taken to minimize the potential for measurement error by employing memory cues, limiting the reference period to a relatively short duration and ensuring that instructions and questions were clear. Two rounds of cognitive interviews were conducted to identify possible sources of confusion in the questionnaire. One especially notable finding from these interviews was that anglers were very eager to report fishing activity and were frustrated by questionnaire versions that did not include time periods that coincided with their fishing activity.

Results from cognitive interviews resulted in several questionnaire versions that were tested over the course of multiple reference waves during 2013 and 2014. Two versions, Q1 and Q2 (Figure

1), are especially relevant for an evaluation of measurement error. Version Q1 was administered during wave 2, 2013 and asked respondents to report the number of shore and private boat fishing trips taken during a single period - the two month reference wave (March/April, 2013 in this case). In contrast, Version Q2 asked respondents to report shore and boat trips for four discrete periods during the prior year, starting with the most recent two months, then progressing backwards in two-month, and then four-month increments. In this version, the reference period of interest - the most recent two months - was bounded by the additional time periods. Version Q2 was administered from wave 2, 2013 through wave 3, 2014. The concurrent administration of the two versions allowed for direct comparisons of survey measures between the questionnaires (i.e between bounded and unbounded versions), while the administration of Version Q2 for successive waves provides an opportunity to evaluate the effect of the recall period (i.e. length of time between a behavior and reporting of the behavior) and question sequence (e.g. first question vs. second question) on reporting for fixed, two-month reference periods.

Our evaluation of recall error also included empirical studies. One study (Andrews et al. 2018) tested a shorter, one-month reference period, as well as a bounding technique to reduce recall error. This study demonstrated that estimates derived from one-month reference waves were not significantly different from “standard” FES estimates, suggesting that FES estimates are not susceptible to omission error - respondents did not seem to forget fishing trips that occurred during the longer two-month reference period. In addition, the study demonstrated that bounding the reference month against the prior month resulted in lower estimated fishing activity than asking only about the reference month. The bounding design - asking respondents to separately report fishing activity for two successive months - also allowed us to examine the effect of recall period and question sequence on reported fishing activity for fixed one-month reference periods.

Finally, we administered an experiment to evaluate the effect of question sequence on reporting of fishing activity. The standard FES questionnaire first asks about shore fishing activities then about boat fishing activities. Within each fishing mode, respondents are asked to report the number of fishing days during the prior two months followed by the number of fishing days during the prior twelve months, inclusive of the reference wave. During data processing, we commonly encounter surveys in which the reported number of two-month fishing days is greater than the number of 12-month fishing days. In 2020, approximately 16% of anglers reported in this illogical manner, and the prevalence of the error was greater for shore fishing (15.7% of anglers who reported shore fishing) than boat fishing (9.3% of anglers who reported boat fishing). Based upon these error rates, as well as the observation from cognitive interviews that anglers are eager to report fishing activity, we hypothesized that FES respondents were telescoping fishing trips into the first fishing question that was asked - days fished from the shore during the previous 2 months. The experiment tested this hypothesis by administering four versions of the FES questionnaire that randomized the order of fishing mode, as well as the order of the two-month and 12-month fishing days questions.

This report describes and summarizes the methods and results from analytical and experimental studies we conducted to evaluate recall error in the FES and offers recommendations for reducing recall error in future survey administrations.

## Methods

### *FES Questionnaire Development Study*

The FES design was tested during 2012-2014 in Massachusetts, New York, North Carolina and Florida (Andrews et al. 2014). During this period, several versions of the questionnaire were tested with a goal of maximizing response from the household population (i.e. minimizing nonresponse bias) and minimizing the potential for measurement errors. Two questionnaire versions, Q1 and Q2 were tested during wave 2, 2013. The primary measures of interest for each questionnaire were the number of shore and private boat fishing trips taken during the previous two months (two-month reference wave). The manner in which this information was collected varied between questionnaires. In Q1 respondents were asked to report the number of fishing days for a single two-month reference wave, while Q2 asked respondents to report the number of fishing days for four discrete periods, beginning with the two-month reference wave (Figure 1). We compared the estimated proportion of households that reported fishing during the two-month reference wave (fishing prevalence) for each fishing mode across questionnaires to determine if the presence/absence of additional reporting periods impacted responses to the primary survey measure (fishing during the reference wave). All analyses were performed using SAS survey procedures in SAS Version 9.4. We used a Chi-Square test of independence to measure the association between the questionnaire version and fishing prevalence.

Figure 1. Questionnaire versions Q1 (left) and Q2 (right) that were administered during FES testing in 2013-2014.

**11. In the past 2 months, between March 1 and April 30, 2013, on how many days did this household member go recreational saltwater fishing in North Carolina from:**

**a. The shore – include docks, bridges, causeways, beaches, banks or any other shore-based structure or area.**

☐☐ **Days fished from shore** *Enter "0" if none.*

**b. A boat – include a private or rental boat that returned to shore in North Carolina. Do not include charter boats - rental or commercial boats that include a captain or crew who help locate and catch fish.**

☐☐ **Days fished from a private boat** *Enter "0" if none.*

**11. For each time period below, on how many days did this person go recreational saltwater fishing in Maryland from:**

**a. The shore – include docks, bridges, causeways, beaches, banks or any other shore-based structure or area. Enter "0" if none.**

☐☐ Days in March and April, 2013

☐☐ Days in Jan. and Feb., 2013

☐☐ Days in Sept., Oct., Nov., Dec., 2012

☐☐ Days in May, June, July, Aug. ,2012

**b. A boat – include a private or rental boat that returned to shore in Maryland. Do not include charter boats - rental or commercial boats that include a captain or crew who help locate and catch fish. Enter "0" if none.**

☐☐ Days in March and April, 2013

☐☐ Days in Jan. and Feb., 2013

☐☐ Days in Sept., Oct., Nov., Dec., 2012

☐☐ Days in May, June, July, Aug., 2012

The inclusion of four reference periods in Q2 allowed us to evaluate the effect of recall length (2 months vs. 4 months) and question sequence (first question vs. second question) on estimated fishing prevalence. For each fishing mode, this questionnaire version asked about fishing activity during the most recent two-month period followed by the prior two-month period (Figure 1). For successive waves, this provided two independent estimates for a fixed reference period, but with different recall lengths and a different position in the question sequence (Figure 2). For example, consider wave 4, 2013, which is shaded blue in Figure 2. The survey administered at the conclusion of wave 4 asked respondents to report the number of fishing days during wave 4 (July-August) followed by the number of fishing days during wave 3 (May-June). In this case, wave 4 was the first question, and the maximum recall length (time between fishing and reporting) for the wave was two months. For the following wave (wave 5, 2013), the questionnaire first asked about fishing during wave 5 then about fishing during wave 4. In this case, wave 4 was the second question, and the maximum recall period was four months. We used a Chi-Square test of independence to measure the association between the recall period/question order and fishing prevalence.

Figure 2. Reference waves for administration of FES Version Q2. Shading identifies overlapping reference periods that vary in the question sequence and recall length. The cross-hatched cells represent the approximate data collection period for each wave.

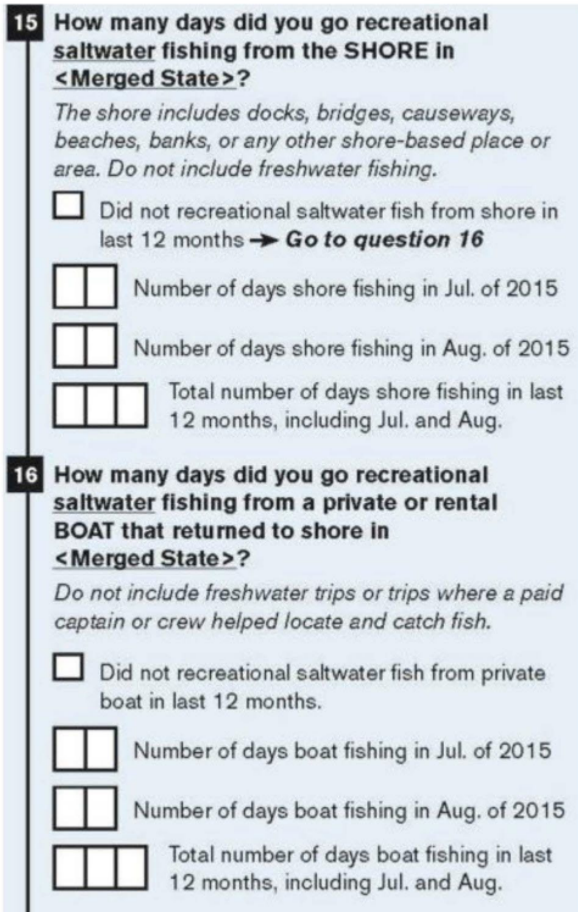
| Year | 2012 | | | | | | 2013 | | | | | | | | | | | | 2014 | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Month | 7 | 8 | 9 | 10 | 11 | 12 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| Survey Waves | Wave 4 - Wave 5 | | | | Wave 6 - Wave 1 | | | | Wave 2 | | Wave 3 | | | | | | | | | | | | | | | |
| | | | Wave 5 - Wave 6 | | | | Wave 1 - Wave 2 | | | | Wave 3 | | Wave 4 | | | | | | | | | | | | | |
| | | | | | Wave 6 - Wave 1 | | | | Wave 2 - Wave 3 | | | | Wave 4 | | Wave 5 | | | | | | | | | | | |
| | | | | | | | Wave 1 - Wave 2 | | | | Wave 3 - Wave 4 | | | | Wave 5 | | Wave 6 | | | | | | | | | |
| | | | | | | | | | Wave 2 - Wave 3 | | | | Wave 4 - Wave 5 | | | | Wave 6 | | Wave 1 | | | | | | | |
| | | | | | | | | | | | Wave 3 - Wave 4 | | | | Wave 5 - Wave 6 | | | | Wave 1 | | Wave 2 | | | | | |
| | | | | | | | | | | | | | Wave 4 - Wave 5 | | | | Wave 6 - Wave 1 | | | | Wave 2 | | Wave 3 | | | |

Version Q2 was administered from wave three, 2013 through wave three, 2014, which provided comparisons for six reference waves (Figure 2). The questionnaire was administered in MA, NY, NC and FL during waves 3-6, 2013 and only NC during waves 1-3, 2014.

*One-Month Wave Study*

In 2015 we completed a pilot study to evaluate the effect of the length of the survey reference period on reporting of recreational fishing activity (Andrews et al 2018). The study was administered from July through December, 2015 with surveys mailed at the end of each month. The study suggested that a one-month reference period, in an unbounded design, is more susceptible to telescoping error than a two-month period, and that bounding a one-month reference period against the prior month reduced telescoping error. Similar to the FES Questionnaire Development Study, the design of the bounded questionnaire asked about fishing activity for successive reference periods. Unlike the prior study, the reporting periods were presented in chronological order - the earlier month was presented first, and the more recent month was presented second (Figure 3). Once again, this design (Figure 4) provided two independent measures for each month that varied in question sequence and recall length. We used a Chi-Square test of independence to measure the association between the recall period/question order and fishing prevalence.

Figure 3. Questionnaires tested in 2015 for the One -Month Wave Study. In the version on the left, the reference month (Aug.) was bounded against the prior month. The version on the right did not include the bounding month.


15 **How many days did you go recreational saltwater fishing from the SHORE in <Merged State>?**

*The shore includes docks, bridges, causeways, beaches, banks, or any other shore-based place or area. Do not include freshwater fishing.*

☐ Did not recreational saltwater fish from shore in last 12 months → ***Go to question 16***

☐☐ Number of days shore fishing in Jul. of 2015

☐☐ Number of days shore fishing in Aug. of 2015

☐☐☐ Total number of days shore fishing in last 12 months, including Jul. and Aug.

16 **How many days did you go recreational saltwater fishing from a private or rental BOAT that returned to shore in <Merged State>?**

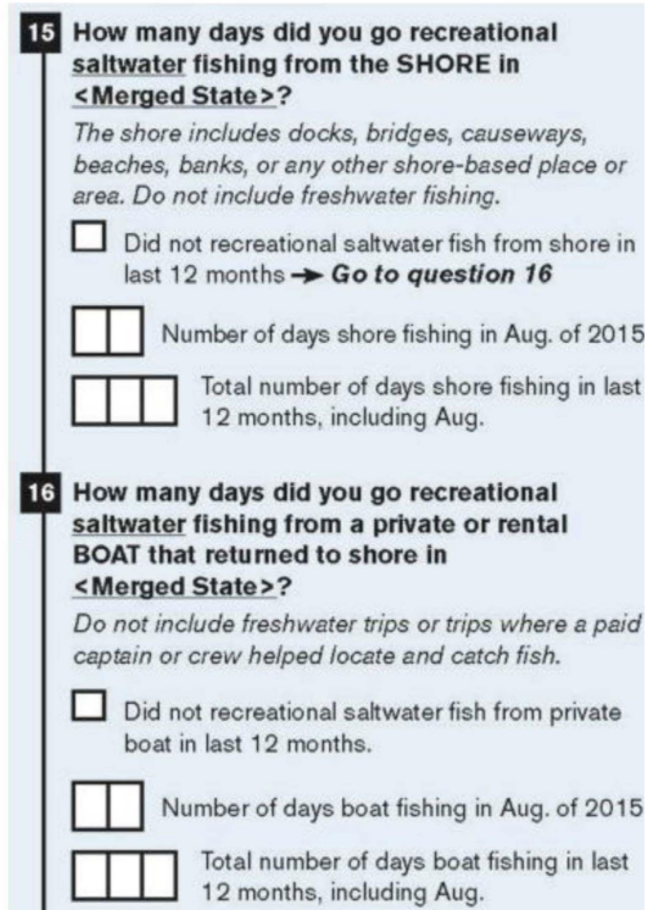
*Do not include freshwater trips or trips where a paid captain or crew helped locate and catch fish.*

☐ Did not recreational saltwater fish from private boat in last 12 months.

☐☐ Number of days boat fishing in Jul. of 2015

☐☐ Number of days boat fishing in Aug. of 2015

☐☐☐ Total number of days boat fishing in last 12 months, including Jul. and Aug.


15 **How many days did you go recreational saltwater fishing from the SHORE in <Merged State>?**

*The shore includes docks, bridges, causeways, beaches, banks, or any other shore-based place or area. Do not include freshwater fishing.*

☐ Did not recreational saltwater fish from shore in last 12 months → ***Go to question 16***

☐☐ Number of days shore fishing in Aug. of 2015

☐☐☐ Total number of days shore fishing in last 12 months, including Aug.

16 **How many days did you go recreational saltwater fishing from a private or rental BOAT that returned to shore in <Merged State>?**

*Do not include freshwater trips or trips where a paid captain or crew helped locate and catch fish.*

☐ Did not recreational saltwater fish from private boat in last 12 months.

☐☐ Number of days boat fishing in Aug. of 2015

☐☐☐ Total number of days boat fishing in last 12 months, including Aug.

Figure 4. One-month reference period study design. Surveys were administered during a two-month period at the conclusion of each reference month (hatched boxes). The questionnaire asked respondents to report fishing activity for two successive months, which were presented in chronological order. The numbers in the boxes represent the question sequence; 1=first question, 2=second question. For successive, monthly survey administrations, a given month (e.g. July) would be the second question in the first survey administration and the first question in the next survey administration. Months adjacent to the data collection periods have a maximum 1-month recall length, while the prior months have a maximum 2-month recall period. Color shading identifies overlapping reference months for successive survey administrations.

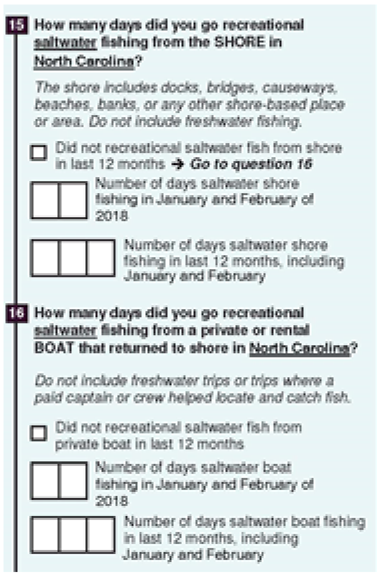
| Month | Jun | July | Aug | Sep | Oct | Nov | Dec | Jan | Feb |
|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | | | | | | | |
| Question Order | | 1 | 2 | | | | | | |
| | | | 1 | 2 | | | | | |
| | | | | 1 | 2 | | | | |
| | | | | | 1 | 2 | | | |
| | | | | | | 1 | 2 | | |

*Question Sequence Experiment*

To empirically quantify the effect of question sequence on reporting of recreational fishing activity, we designed experimental questionnaires that randomized the sequence of fishing mode, as well as the sequence of the two-month and 12-month fishing days questions (Figure 5). The study was conducted in all Atlantic and Gulf of Mexico states covered by the FES (ME-MS) during waves 4-6, 2019. All sampling, data collection, data processing and weighting protocols were identical to those for the production FES (Papacostas and Foster 2018). For each wave, a sample was selected from the FES sample frame and randomly allocated into “base” FES and experimental treatments. Sample sizes and response rates are provided in Table 1.

Figure 5. Questionnaire versions tested in the Question Sequence Experiment during 2019. The questionnaire versions varied in the sequence of fishing mode and two-month and 12-month fishing activity questions.

FES: SH:PR/2:12



**15** **How many days did you go recreational saltwater fishing from the SHORE in North Carolina?**

*The shore includes docks, bridges, causeways, beaches, banks, or any other shore-based place or area. Do not include freshwater fishing.*

☐ Did not recreational saltwater fish from shore in last 12 months → ***Go to question 16***

Number of days saltwater shore fishing in January and February of 2018

Number of days saltwater shore fishing in last 12 months, including January and February

**16** **How many days did you go recreational saltwater fishing from a private or rental BOAT that returned to shore in North Carolina?**

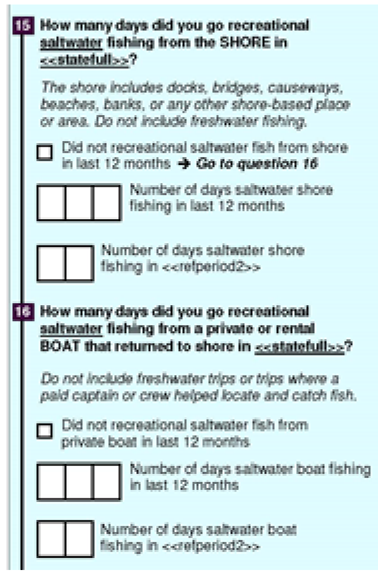
*Do not include freshwater trips or trips where a paid captain or crew helped locate and catch fish.*

☐ Did not recreational saltwater fish from private boat in last 12 months

Number of days saltwater boat fishing in January and February of 2018

Number of days saltwater boat fishing in last 12 months, including January and February

EXP1: SH:PR/12:2



**15** **How many days did you go recreational saltwater fishing from the SHORE in <<statefull>>?**

*The shore includes docks, bridges, causeways, beaches, banks, or any other shore-based place or area. Do not include freshwater fishing.*

☐ Did not recreational saltwater fish from shore in last 12 months → ***Go to question 16***

Number of days saltwater shore fishing in last 12 months

Number of days saltwater shore fishing in <<refperiod2>>

**16** **How many days did you go recreational saltwater fishing from a private or rental BOAT that returned to shore in <<statefull>>?**

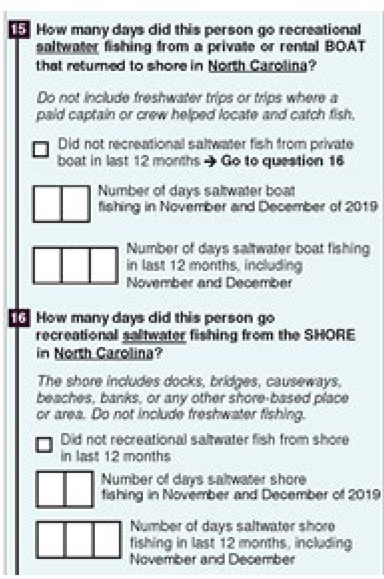
*Do not include freshwater trips or trips where a paid captain or crew helped locate and catch fish.*

☐ Did not recreational saltwater fish from private boat in last 12 months

Number of days saltwater boat fishing in last 12 months

Number of days saltwater boat fishing in <<refperiod2>>

EXP2: PR:SH/2:12



**15** **How many days did this person go recreational saltwater fishing from a private or rental BOAT that returned to shore in North Carolina?**

*Do not include freshwater trips or trips where a paid captain or crew helped locate and catch fish.*

☐ Did not recreational saltwater fish from private boat in last 12 months → **Go to question 16**

Number of days saltwater boat fishing in November and December of 2019

Number of days saltwater boat fishing in last 12 months, including November and December

**16** **How many days did this person go recreational saltwater fishing from the SHORE in North Carolina?**

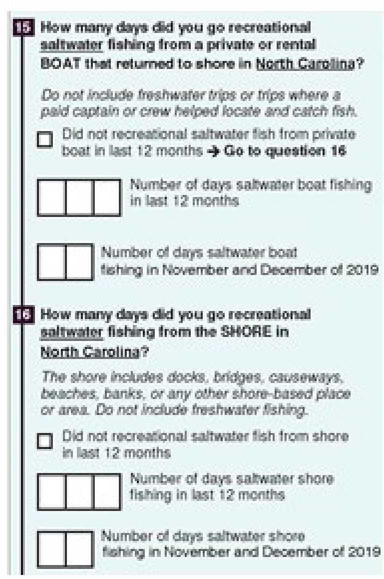
*The shore includes docks, bridges, causeways, beaches, banks, or any other shore-based place or area. Do not include freshwater fishing.*

☐ Did not recreational saltwater fish from shore in last 12 months

Number of days saltwater shore fishing in November and December of 2019

Number of days saltwater shore fishing in last 12 months, including November and December

EXP3: PR:SH/12:2



**15** **How many days did you go recreational saltwater fishing from a private or rental BOAT that returned to shore in North Carolina?**

*Do not include freshwater trips or trips where a paid captain or crew helped locate and catch fish.*

☐ Did not recreational saltwater fish from private boat in last 12 months → **Go to question 16**

Number of days saltwater boat fishing in last 12 months

Number of days saltwater boat fishing in November and December of 2019

**16** **How many days did you go recreational saltwater fishing from the SHORE in North Carolina?**

*The shore includes docks, bridges, causeways, beaches, banks, or any other shore-based place or area. Do not include freshwater fishing.*

☐ Did not recreational saltwater fish from shore in last 12 months

Number of days saltwater shore fishing in last 12 months

Number of days saltwater shore fishing in November and December of 2019

Table 1. Question sequence experiment sample sizes, completed interviews and response rates (AAPOR RR1) by questionnaire

| | FES SH:PR/2:12 | | | EXP1 SH:PR/12:2 | | | EXP2 PR:SH/2:12 | | | EXP3 PR:SH/12:2 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Wave | Sampled Addresses | Complete (n) | Complete (%) | Survey Mailings | Complete (n) | Complete (%) | Survey Mailings | Complete (n) | Complete (%) | Survey Mailings | Complete (n) | Complete (%) |
| 4 | 40,266 | 12,250 | 30.4 | 11,482 | 3,535 | 30.8 | 11,483 | 3,501 | 30.5 | 11,482 | 3,456 | 30.1 |
| 5 | 53,228 | 16,150 | 30.3 | 15,176 | 4,614 | 30.4 | 15,178 | 4,621 | 30.4 | 15,179 | 4,533 | 29.9 |
| 6 | 52,412 | 16,351 | 31.2 | 14,948 | 4,630 | 31 | 14,947 | 4,533 | 30.3 | 14,946 | 4,680 | 31.3 |
| Total | 145,906 | 44,751 | 30.7 | 41,606 | 12,779 | 30.7 | 41,608 | 12,655 | 30.4 | 41,607 | 12,669 | 30.4 |

The objectives of the study were to compare measures of fishing activity among survey questionnaires and evaluate the frequency of illogical responses in which the number of wave (two-month) fishing days exceeded the number of annual (12-month fishing days). We used logistic regression to compare fishing propensity (prevalence) among questionnaire versions.

**Results**

*FES Questionnaire Development Study*

Table 2 compares estimated shore and private boat fishing prevalence between Q1 and Q2 overall and by state. Questionnaire version 2, which bounded the 2-month wave against additional time-periods, resulted in lower estimates than version 1 in nine out of ten comparisons; the sole exception was for shore fishing in Florida. With the exception of Florida shore fishing, version 2 estimates were 13.27% - 86.94% lower than version 1 estimates. Chi square tests of independence identified significant associations between questionnaire version and prevalence for shore fishing in MA ($p=0.0208$) and for private boat fishing overall ($p=0.0002$) and in MA ($p=0.0054$), NY ($p<0.0001$).

Table 2. Estimated shore and private boat fishing prevalence overall and by state for unbounded and bounded questionnaire versions. Relative difference is expressed as a percent and represents the difference between Q1 and Q2 with respect to Q1.

| | Q1 (unbounded) | Q2 (bounded) | Relative difference | |
|---|---|---|---|---|
| Shore prevalence | % (SE) | % (SE) | % | p-value |
| Overall | 9.42 (0.62) | 9.27 (0.62) | 1.59 | 0.8659 |
| FL | 16.20 (1.31) | 19.34 (1.43) | -19.40 | 0.1047 |
| MA | 6.61 (1.15) | 3.40 (0.81) | 48.65 | 0.0208 |
| NY | 4.76 (0.87) | 3.13 (0.77) | 34.20 | 0.1672 |
| NC | 7.37 (1.35) | 5.86 (1.19) | 20.50 | 0.4009 |
| Boat prevalence | | | | |
| Overall | 7.54 (0.55) | 4.92 (0.43) | 34.75 | 0.0002 |
| FL | 13.39 (1.18) | 11.62 (1.12) | 13.27 | 0.2734 |
| MA | 4.17 (0.81) | 1.38 (0.51) | 67.00 | 0.0054 |
| NY | 4.89 (0.91) | 0.64 (0.37) | 86.94 | <0.0001 |
| NC | 3.71 (0.92) | 2.77 (0.60) | 25.39 | 0.3721 |

Table 3 compares shore and private boat fishing prevalence by recall length, which were estimated from Questionnaire Version Q2, overall and for each reference wave. For shore fishing, estimates derived from the shorter recall period, which is also the first fishing question asked of each respondent, are nominally larger than estimates derived from the longer recall period for 12 of the 16 comparisons. The difference in shore prevalence between recall periods is significant overall (p=0.0346). Results are less systematic for estimated boat fishing prevalence - estimates derived from the 2-month recall period are larger in 9 of 16 comparisons, and none of the differences are significant.

Table 3. Estimated shore and boat fishing prevalence by recall length, overall and by state and reference wave. Relative difference is expressed as a percent and represents the difference between estimates relative to the estimate derived from the 2-month recall period.

| | | Shore | | | | Boat | | | |
|---|---|---|---|---|---|---|---|---|---|
| State | Wave | 2-mo recall % (SE) | 4-mo recall % (SE) | Relative difference | *p*-value | 2-mo recall % (SE) | 4-mo recall % (SE) | Relative difference | *p*-value |
| FL | May/Jun | 21.63 (1.11) | 18.73 (1.32) | 13.42 | 0.0970 | 11.6 (0.81) | 11.04 (0.99) | 4.84 | 0.6619 |
| FL | Jul/Aug | 20.47 (1.38) | 21.86 (1.62) | -6.75 | 0.5148 | 10.98 (0.99) | 14.50 (1.34) | -32.04 | 0.0311 |
| FL | Sep/Oct | 19.21 (1.55) | 16.51 (1.33) | 14.03 | 0.1853 | 12.04 (1.25) | 9.71 (1.00) | 19.40 | 0.1401 |
| MA | May/Jun | 10.83 (0.91) | 9.94 (1.19) | 8.23 | 0.5557 | 6.03 (0.65) | 5.80 (0.94) | 3.83 | 0.8412 |
| MA | Jul/Aug | 15.86 (1.50) | 15.58 (1.62) | 1.71 | 0.9025 | 9.39 (1.13) | 11.07 (1.30) | -17.85 | 0.3274 |
| MA | Sep/Oct | 7.33 (1.17) | 8.72 (1.23) | -19.05 | 0.4112 | 5.15 (0.89) | 5.37 (1.03) | -4.24 | 0.8722 |
| NY | May/Jun | 8.86 (0.77) | 7.35 (1.11) | 17.10 | 0.2785 | 5.81 (0.64) | 4.64 (0.88) | 20.22 | 0.3010 |
| NY | Jul/Aug | 12.93 (1.46) | 12.62 (1.51) | 2.41 | 0.8824 | 8.25 (1.18) | 10.87 (1.47) | -31.81 | 0.1593 |
| NY | Sep/Oct | 7.06 (1.13) | 6.21 (0.89) | 12.10 | 0.5489 | 6.70 (1.16) | 4.93 (0.82) | 26.43 | 0.2015 |
| NC | May/Jun | 11.5 (0.91) | 11.32 (1.21) | 1.54 | 0.9077 | 5.87 (0.61) | 4.89 (0.71) | 16.80 | 0.3017 |
| NC | Jul/Aug | 13.38 (1.28) | 12.43 (1.21) | 7.07 | 0.5904 | 5.93 (0.77) | 7.82 (1.04) | -31.80 | 0.1350 |
| NC | Sep/Oct | 10.71 (1.19) | 11.19 (1.45) | -4.50 | 0.7967 | 6.13 (0.93) | 5.18 (0.90) | 15.49 | 0.4643 |
| NC | Nov/Dec | 6.42 (1.09) | 4.57 (1.12) | 28.84 | 0.2488 | 2.69 (0.61) | 2.86 (0.92) | -6.47 | 0.8736 |
| NC | Jan/Feb | 3.06 (1.08) | 3.59 (1.30) | -17.28 | 0.7528 | 1.14 (0.26) | 0.84 (0.18) | 26.95 | 0.3153 |
| NC | Mar/Apr | 8.77 (1.88) | 5.30 (0.83) | 39.59 | 0.0573 | 3.20 (1.09) | 2.36 (0.54) | 26.30 | 0.4548 |
| Overall | | 13.1 (0.34) | 12.04 (0.36) | 8.09 | 0.0346 | 7.53 (0.25) | 7.57 (0.29) | -0.53 | 0.9054 |

*One-Month Wave Study*

Table 4 compares estimated prevalence by recall length/question sequence for each month of the one-month wave study. In 12 of 12 comparisons, estimated prevalence was nominally lower when the reference month was presented second in the question sequence, which also coincided with a shorter recall length (one month prior to survey administration vs. two months). For shore fishing, estimates were 7.3-49.7% lower when the reference month was presented second in the sequence, and differences between estimates were significant overall (p=0.0002) and for July (p=0.0262), August (p=0.0286) and November (p=0.0409). For boat fishing, estimates were 24.9-47.4% lower when the reference month was presented second, and differences were significant overall (p=0.0002) and for August (p=0.0025) and November (p=0.0264).

Table 4. Estimated shore and private boat fishing prevalence overall and by month for questionnaire versions in which the reference month was presented first with a longer recall (Month 1) and second with a shorter recall (Month 2). Relative difference is expressed as a percent and represents the difference between estimates relative to the estimate derived from the second month in the sequence (month 2).

| Measure | Month 1 | Month 2 | Relative difference | *p*-value |
|---|---|---|---|---|
| Shore prevalence | | | | |
| Overall | 9.26 | 5.97 | 35.53 | 0.0002 |
| July | 7.82 | 7.25 | 7.30 | 0.7516 |
| August | 12.05 | 6.06 | 49.68 | 0.0025 |
| September | 10.66 | 7.00 | 34.33 | 0.104 |
| October | 7.01 | 4.97 | 29.15 | 0.2448 |
| November | 8.75 | 4.60 | 47.48 | 0.0002 |
| Boat prevalence | | | | |
| Overall | 6.35 | 3.85 | 39.37 | 0.0002 |
| July | 6.46 | 3.44 | 46.73 | 0.0262 |
| August | 7.62 | 4.43 | 41.92 | 0.0386 |
| September | 6.06 | 4.20 | 30.69 | 0.2453 |
| October | 4.67 | 3.51 | 24.89 | 0.4177 |
| November | 6.95 | 3.66 | 47.39 | 0.0409 |

*Question Sequence Experiment*

Figure 6 compares two-month shore and private boat fishing prevalence among the questionnaire versions included in the question sequence experiment. For shore fishing, both the mode sequence (p=0.0034) and sequence of the two-month and 12-month fishing questions (p=0.0002) were significant predictors of reporting fishing activity during the 2-month reference wave. For shore fishing, estimated prevalence was 16% lower when boat fishing preceded shore fishing and nearly 20% lower when the 12-month question preceded the 2-month question. For boat fishing, only the sequence of the two-month and 12-month questions was significant for predicting fishing activity (p<0.0001). Estimated boat fishing prevalence was nearly 30% lower when the 12-month question preceded the 2-month question. The interaction between fishing mode sequence and two-month / 12-month sequence was not significant for either shore or boat fishing. However, pairwise comparisons among questionnaires demonstrated that estimated prevalence was highest for both fishing modes when the mode was presented first and the two-month fishing question preceded the 12-month question (FES for shore and EXP2 for boat). Within the two-month/12-month sequence, shore and boat prevalence were both significantly lower when the mode was presented second (EXP2 for shore and FES for boat).

Figure 6. Estimated prevalence of fishing activity during the two-month reference wave by fishing mode for the questionnaire versions tested in the Question Sequence Experiment conducted in 2019. Bars represent the proportion of households that reported fishing during the reference wave (+/- 95% CI). Within a fishing mode, bars with different letters are significantly different (*p*<0.05).

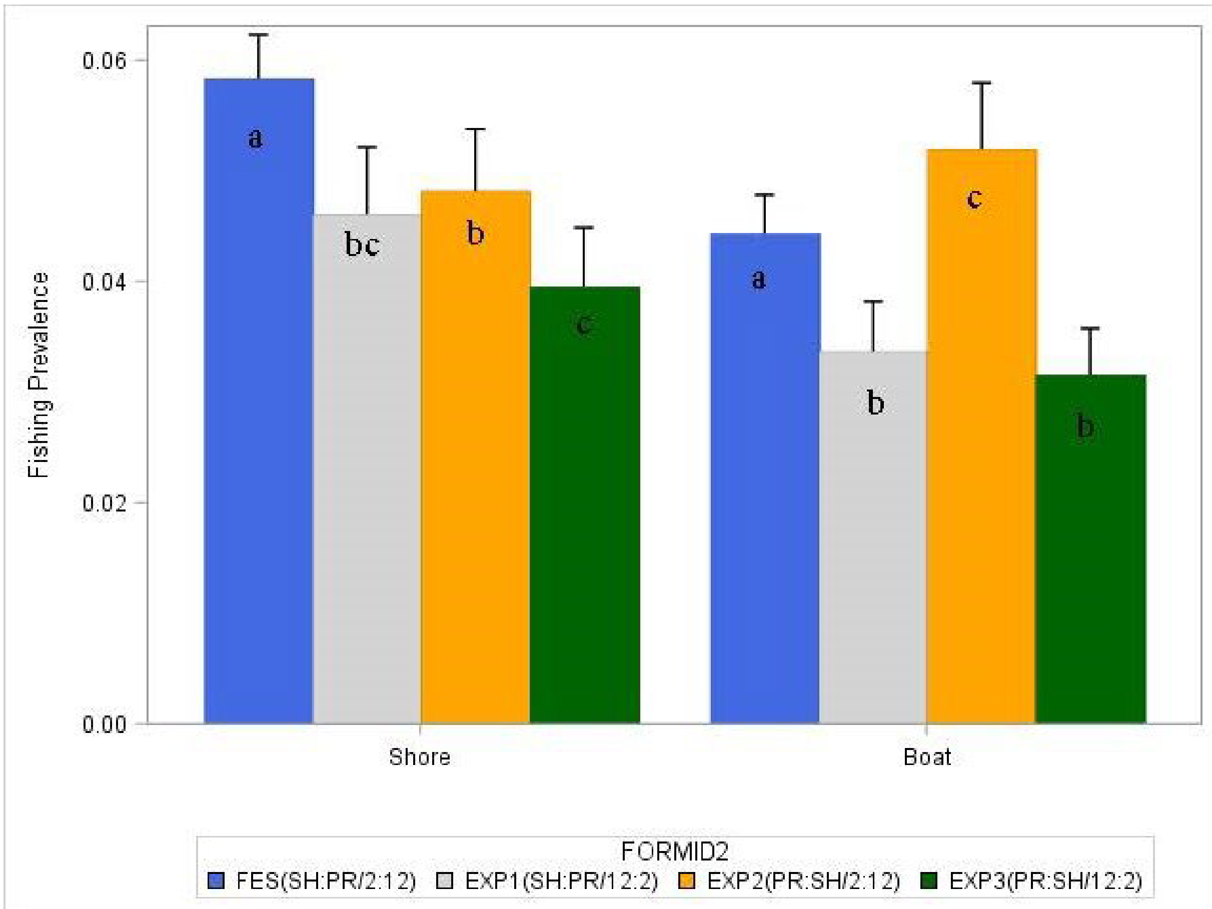



Figure 7 shows the ratios of experimental to FES total effort estimates by fishing mode across all states and waves in which the study was conducted. The patterns are consistent to what we observed for prevalence - estimates are lower (i.e. ratios are lower) for the mode that is listed second on the questionnaire and when the 12-month fishing days question precedes the two-month fishing days question. For shore fishing, the mean ratios across all states and waves were 0.71, 0.92 and 0.62 for EXP1, EXP2 and EXP3, respectively. For private boat fishing, the mean ratios were 0.72, 1.24 and 0.70 for the respective questionnaires.

Figure 7. Box plots of total effort ratios across all states and waves. Ratios are of experimental estimates to FES estimates. Ratios are less than 1.0 when FES estimates are larger than experimental estimates. Within each box the circle and horizontal line represent the mean and median ratios, respectively.

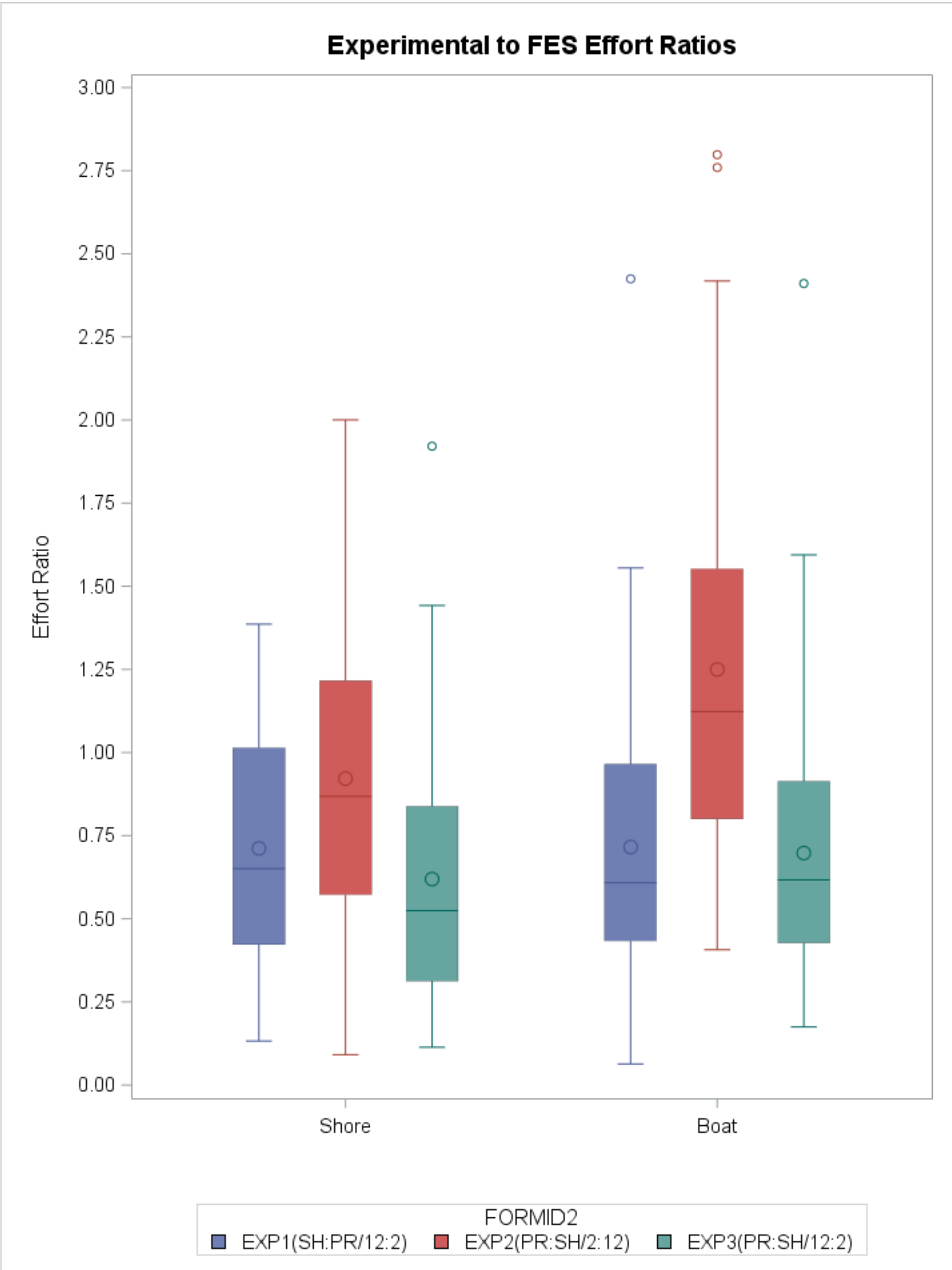


Table 5 shows the percentage of responding anglers who reported more days fishing during the two-month wave than the previous 12 months, which included the two-month wave. For both fishing modes, error rates were nearly 80% lower for the two questionnaire versions that asked about fishing during the prior year before asking about fishing during the two-month wave.

Table 5. Percent of respondents who reported more fishing days during the reference wave than the previous 12 months by questionnaire version.

| Questionnaire | Shore trips | Boat trips |
|---|---|---|
| FES | 6.92 | 5.97 |
| EXP1 | 1.6 | 1.12 |
| EXP2 | 7.9 | 5.22 |
| EXP3 | 1.66 | 1.28 |

**Discussion**

Cognitive interviews during FES development suggested that anglers were eager to report fishing activity, even when the activity was not within the intended time-period of the survey. The results presented here are consistent with that observation and support the hypothesis that bounding techniques, including offering multiple reporting periods and re-ordering the sequence of fishing questions, may reduce recall error.

Bounded recall is a common technique used to reduce telescoping error in retrospective surveys (Sudman et al. 1984). Andrews et al. (2018) suggested that bounding the reference period against a second, earlier period may reduce telescoping error in recreational fishing surveys with one-month reference periods. The effect of bounding can be seen in our comparison of historical FES data (Table 2). In nine of the 10 comparisons, bounded estimates were as much as 87% lower than unbounded estimates.

Previous studies (Sudman et al. 1984, Loftus et al. 1990, Andrews et al. 2018) found that telescoping was reduced when the reference period was bounded by an earlier period or a longer period that included the reference period. The survey questions in these studies were sequenced such that the bounding period was introduced before the desired reference period - respondents were asked about a behavior during the bounding period and then asked about the behavior during the reference period. In our initial comparison between bounded and unbounded designs, described above, the sequence was reversed - anglers were asked to report the number of fishing days during the reference period before reporting for multiple bounding periods.

Using the same data, we evaluated the effect of question sequence, as well as the recall length, on reporting and found that fishing prevalence estimates were larger when the reference period was displayed on the questionnaire prior to the bounding period. This could suggest that anglers telescoped trips into the first reporting period in the sequence, which would be consistent with the observation from cognitive interviews that anglers are eager to report fishing activities. However, the lower estimates also coincided with a longer recall period (e.g. longer period of time between fishing and reporting), which suggests recall decay/omission error as a potential source of the observed differences between estimates - i.e. anglers may have forgotten about trips further back in time.

Results were more pronounced when the reference period was limited to a single month. In this study, fishing prevalence estimates for a given reference month were lower when the month was displayed second in the two-month sequence than when it was displayed first. In this case, the smaller estimates coincided with a shorter recall period, refuting the suggestion that recall decay contributed to the differences.

The current FES questionnaire includes a two-month reference period followed by a 12-month bounding period. This design was initially considered because memory retrieval should be easier for a shorter time-period than a longer one - i.e. we wanted to ask the easy question first (Schwartz et al. 2008). However, results from the question sequence experiment provide evidence that this design is not the most effective for reducing telescoping error- prevalence estimates were as much as 32% and 39% lower for shore and private boat fishing, respectively, when the 12-month bounding period preceded the two-month reference period. Similar differences among questionnaire treatments were observed when we compared estimates of total effort. Telescoping using the two-month / 12-month sequence seems to be reduced somewhat for the second mode in the sequence (boat mode in the FES), but even these estimates are larger than estimates derived from the 12-month/two-month sequence.

This discussion assumes that lower estimates are more accurate than higher estimates - i.e. that telescoping is the primary error in unbounded designs. This assumption is supported by Loftus et al. (1990), who validated survey responses against administrative records and concluded that bounded designs reduce over-reporting of behaviors. Our review of error rates in the question sequence study also supports this assumption as illogical responses (reporting more trips during the two-month reference wave than the 12-month bounding period) were significantly lower when the 12-month period was presented before the two-month reference period.

Loftus et al. (1990) describe several mechanisms by which bounded recall may reduce telescoping and improve the accuracy of self-reports. These include, 1) stimulating and improving memory, 2) satisfying a need to be helpful or to report a socially desirable behavior, and 3) conveying a need for greater detail. Based upon anecdotal information from cognitive interviews, as well as the effect of question sequence on reporting, we suspect that anglers are so eager to report fishing activity, that they do so at the earliest possible opportunity, even if it means providing inaccurate, out-of-scope information. Such a mechanism is similar to satisfying a need to be helpful, but also incorporates the sense of pride and identity that was expressed by anglers during cognitive interviews. Presenting a 12-month fishing question prior to the two-month question provides respondents with an opportunity to identify themselves as anglers that is both more accurate (the probability that a respondent fished during the prior 12 months is greater than the prior 2 months) and likely to reduce telescoping error for the primary survey measure (two-month fishing activity).

One notable distinction between the prior studies and the current analysis is that the previous studies utilized either interviewer-administered surveys or a sequence of self administered questionnaires that did not permit respondents to view the entire questionnaire (Sudman et al. 1984, Loftus 1990). In contrast, the FES is a self-administered questionnaire that respondents can review in its entirety before completing. The fact that telescoping error is most prominent for the first fishing question that is asked suggests that respondents are not carefully reviewing the entire questionnaire prior to completing the survey.

**Conclusions**

Results from these studies suggest that the predominant form of measurement error in the FES is telescoping error; respondents were more likely to report trips from outside the intended reference period than omit trips. The current FES questionnaire design includes bounding questions (12-month shore and boat fishing) that are likely to reduce telescoping error in the FES relative to an unbounded design. However, the question sequence may not be optimal in terms of reducing telescoping error - respondents are unlikely to review the entire questionnaire, including the bounding questions, prior to reporting for the desired reference period. Consequently, the current design is likely overestimating fishing effort. A revised design that presents the 12-month fishing questions before the two-month questions would likely further reduce bias resulting from telescoping error.

**References**

Andrews, R., J.M. Brick and N. Mathiowetz. 2014. Development and testing of recreational fishing effort surveys: Testing a mail survey design. Retrieved from: https://www.st.nmfs.noaa.gov/pims/main/public?method=DOWNLOAD_FR_PDF&record_id=1179.

Andrews, W.R., K.J. Papacostas, and J. Foster. 2018. A Comparison of Recall Error in Recreational Fisheries Surveys with One- and Two-Month Reference Periods. North American Journal of Fisheries Management. 38: 1284-1298.

Chu, A., D. Eisenhower, M. Hay, D. Morganstein, J. Neter, and J. Waksberg. 1992. Measuring the recall error in self-reported fishing and hunting activities. Journal of Official Statistics 8(1):19–39.

Gaskell, G.D., D.B. Wright, and C.A. O'Muircheartaigh. 2000. Telescoping of Landmark Events: Implications for Survey Research. Public Opinion Quarterly 64(1): 77-89.

Loftus, E.F., M.R. Klinger, K.D. Smith, and J. Fielder. 1990. A Tale of Two Questions: Benefits of Asking More than One Question. Public Opinion Quarterly 54:330-345.

Loftus, E. F., and W. Marburger. 1982. Since the eruption of Mt. St. Helens, has anyone beaten you up? Improving the accuracy of retrospective reports with landmark events. Memory and Cognition 11:114-120.

National Academies of Sciences, Engineering, and Medicine. 2017. Review of the Marine Recreational Information Program. Retrieved from https://www.nationalacademies.org/our-work/review-of-the-marine-recreational-information-program.

Neter, J., and J. Waksberg. 1964. A Study of Response Errors in Expenditures Data from Household Interviews. Source Journal of the American Statistical Association 59(305):18–55.

Papacostas, K. and J. Foster. 2018. National Marine Fisheries Service's Marine Recreational Information Program survey design and statistical methods for estimation of recreational fisheries catch and effort. Retrieved from https://www.fisheries.noaa.gov/resource/document/survey-design-and-statistical-methods-estimation-recreational-fisheries-catch-and.

Schwartz, N., B. Knauper, D. Oyserman, C. Stitch. 2008. The Psychology of Asking Questions. In E.D. de Leeuw, J.J. Hox, D.A. Dillman (Eds.). International Handbook of Survey Methodology (pp. 18-34). New York, New York: Routledge.

Sudman, S., and N. M. Bradburn. 1973. Effects of Time and Memory Factors on Response in Surveys Efects of Time and Memory Factors on Response in Surveys. Journal of the American Statistical Association 68206112(344):805–815.

Sudman, S., A. Finn, and L. Lannom. 1984. The Use of Bounded Recall Procedures in Single Interviews. Public Opinion Quarterly 48(2):520.

**Attachment A: Consultant Report**

**Review of Recall Error and Measurement Error Documents**

Jay Breidt, Jean Opsomer and Mike Brick

June 22, 2023

We have reviewed the two documents (A comparison of recall error in recreational fisheries surveys with one and two-month reference periods; Evaluating measurement error in the MRIP Fishing Effort Survey: The effect of question sequence on reporting of fishing activity) that describe measurement error studies conducted regarding the Fishing Effort Survey (FES). Before discussing the specifics of the reports, we wish to recognize NOAA Fisheries and the MRIP team on a thoughtful and consistent program of working to understand and improve the FES over more than a decade. The profound thinking about sources of errors and the commitment of resources to implement studies to examine these error sources and their consequences is outstanding. It is a clear indicator of the value placed on attaining high quality estimates from a statistical system.

We provide detailed comments on each article below, but in general both provide excellent analyses exploring potential sources of measurement error in the FES. The exposition is clear and the rigor is high. The studies describe what is known and what is not known, state hypotheses regarding measurement errors, describe the studies developed and implemented to explore the hypotheses, and report the analysis and recommendations from those studies. We consider these to be scientifically and methodologically sound and found no issues with the work.

Our comments aim to either add context or to suggest further analysis that may enhance the reports for readers. One overarching comment is that the two articles focus entirely on measurement error, but some readers might not understand that nonresponse bias was initially the nonsampling error of greatest concern in fishing effort surveys (the 2006 National Academy Report). Considerable resources were devoted to nonresponse bias, and it might help to acknowledge this (possibly in the first paragraph of the executive summary) when disseminating these reports so that the context is clear.

*A comparison of recall error in recreational fisheries surveys with one and two-month reference periods* by Andrews, Papacostas, and Foster

The article has already been through the academic peer-review process and has been published, which is an important statement about the quality of the article and its material. Since our comments are more editorial than substantive, we understand that providing the published article without making revisions based on our comments might be preferred or dealing with revisions in supplemental materials or an appendix to the published article.

Detailed comments:

1. Figure 1 uses the term 'Mailing #' which is somewhat confusing since it refers to a data collection that has multiple mailings. One suggestion to avoid the term "mailing" is to say "September survey, October survey, etc."? This would also emphasize the fact that the surveys are conducted immediately after the reference period in the survey. Alternatively, just write "survey 1, survey 2, etc."
2. Page 5. Technically, the estimator is not the Horvitz-Thompson (HT) estimator (it is a nonresponse adjusted post-stratified estimator) although it begins as an HT before the subsequent adjustments that are described.
3. Page 5. "…use of an established procedure for trimming the estimated mean square error (see Potter 1990) to minimize the effects of extreme weights on the sampling variance." The Potter procedure trims extreme weights (not the estimated mean square error) to minimize the estimated mean square error.
4. Table A.1 does not show the response rates although the title says it does. We computed some response rates from the data and agree the differences are small.
5. The Figure 2 analysis is not completely clear. Were the differences from the FES compared to each of the treatments separately or combined? We believe they were separated but some clarification might be useful.
6. Last sentences of Discussion. You discuss precision but do not make the point explicitly that a 1-month recall requires a larger sample size (potentially substantially larger) to achieve the same precision level. Maybe it is obvious to everyone?

*Evaluating measurement error in the MRIP Fishing Effort Survey: The effect of question sequence on reporting of fishing activity* by Andrews

The article provides a summary of a series of experiments that shed light on how the questionnaire structure influences the reporting of fishing prevalence. Since the 'truth' is never completely known in these situations, the author makes plausible arguments to show the designs that are likely to lead to reduce measurement errors and estimates that closer to the unknown truth. The contention in the analysis and explicitly in the Discussion is that lower estimates are more accurate. In other settings, particularly in reporting rare and sensitive items, researchers often assume 'more is better.' We believe the author is making the correct assumption for the FES and provides several arguments to that effect, but addressing the alternative in the discussion might help readers understand the rationale more completely. The author also uses the word 'predictor' in discussing findings from chi-square tests. It would be better to say they are associated with prevalence.

Just below Table 1, the author states that ` We used logistic regression to compare fishing propensity (prevalence) among questionnaire versions.' However, logistic regression is not mentioned again, and results are described in terms of chi-squared tests of independence and complex tables of point estimates by state, wave, fishing mode, and questionnaire type. A logistic regression analysis might be considered further for powerful and interpretable analysis of these complex results. For example, a possibly useful alternative to Table 3 would be survey-weighted logistic regression of trips (yes or no) on categorical variables that include state, wave, fishing mode, and recall length, together with various interactions as supported by the data. One could then formally test for any treatment effect of recall length on fishing prevalence estimates, controlling for the other non-treatment factors.

At the end of the discussion, the author suggests that 'the respondents are not carefully reviewing the entire questionnaire prior to completing the survey.' While this is possible, we wonder whether this is the main issue, especially since with a hard-copy questionnaire it is hard to avoid seeing the items in proximity to each other. Another possibility is that the desire to report fishing is strong and the respondent simply uses the first opportunity to do that. The fact that the difference is less substantial when boating is asked first also raises some suspicions that the mechanism may be more complex[1].

The conclusion that the 12-month questions first is preferred seems well-supported by the data. While the results are not as clear, our impression is that asking private boat before shore sequence is giving consistent and reasonable estimates (Figure 7). We think the recommendation to ask private boating first is reasonable unless there are other issues that we have not considered.

*Comments on Executive Summary*

The executive summary accompanying the two documents briefly describes the main results and makes a reasonable suggestion for how to revise the FES questionnaire to reduce the measurement error in the fishing effort estimates. In particular, we agree with its recommendation to switch the order of the 12-month and 2-month recall questions.

[1] Analysis of households with two persons reporting fishing activity might provide more evidence about this hypothesis. If the hypothesis is true, then the second person with reported trips should not show the same patterns as the first angler (since the items are identical for each person). We admit this analysis is not likely to be very enlightening since the desire to report being an angler is also not likely to be as strong for another person in the household.