North American Journal of Fisheries Management

ISSN: 0275-5947 (Print) 1548-8675 (Online) Journal homepage: http://www.tandfonline.com/loi/ujfm20

Concurrently Assessing Survey Mode and Sample Size in Off-Site Angler Surveys

Boyd N. Barrett, Brett van Poorten, Andrew B. Cooper & Wolfgang Haider

North American Journal of Fisheries Management 37:756–767, 2017

ISSN: 0275-5947 print / 1548-8675 online
DOI: https://doi.org/10.1080/02755947.2017.1324543

Check for updates

ARTICLE

Concurrently Assessing Survey Mode and Sample Size in Off-Site Angler Surveys

Boyd N. Barrett*
School of Resource and Environmental Management, Simon Fraser University, 8888 University Drive, Burnaby, British Columbia V5A 1S6, Canada

Brett van Poorten
British Columbia Ministry of Environment, 2202 Main Mall, Vancouver, British Columbia V6T 1Z4, Canada

Andrew B. Cooper and Wolfgang Haider
School of Resource and Environmental Management, Simon Fraser University, 8888 University Drive, Burnaby, British Columbia V5A 1S6, Canada

Abstract
Off-site angler surveys are commonly administered via two or more survey modes in the form of a mixed-mode survey. Mixed-mode surveys allow survey administrators to attain the benefits inherent to different survey modes, reduce total survey error, and control survey cost. However, these benefits can only be simultaneously attained after undertaking sample size planning. Sample size planning is a trade-off analysis wherein a researcher concurrently assesses survey administration cost, the accuracy and precision of estimates, the magnitude and direction of biases, and variance of the test statistic to determine an optimal sample size. We used data from an off-site angler survey administered to anglers targeting White Sturgeon *Acipenser transmontanus* to illustrate a systematic approach to sample size planning. Our survey design included a mixed-mode design with three survey modes (e-mail, mail, and telephone) and a two-phase sampling design that had a first contact and a follow-up contact with a subsample of nonrespondents. Sample size planning was undertaken in the form of a sensitivity analysis wherein four survey design alternatives were simulated and assessed based on four criteria (i.e., bias, precision, accuracy, and cost). We also incorporated tests for nonresponse bias and survey mode effect. We found that (1) response rates were lower for e-mail surveys (22%) than for mail surveys and telephone surveys (39–44%); (2) nonresponse bias did not have a substantial effect on survey estimates from the mixed-mode design; and (3) estimates (total effort and total catch) from the mail and e-mail survey modes were significantly different, indicating a survey mode effect. The high variability of anglers' annual catch made survey estimates highly imprecise at lower sample sizes. The level of acceptable error varies for each study. Therefore, a systematic approach to sample size planning is necessary to determine the point where acceptable error is reached while considering multiple survey design alternatives.

The collection of reliable effort and catch estimates from recreational fishers is important because these data are used as indices of abundance for many valuable species (Maunder and Punt 2004). However, unlike commercial fisheries, recreational fisheries often do not require mandatory reporting (Eero et al. 2014; Hartill and Edwards 2015). As such, recreational fishery managers rely on the voluntary participation of anglers in on-site (e.g., creel survey) and off-site (e.g., telephone, mail, and internet) surveys to gather effort and catch data (Roach et al. 1999). Decisions on which survey method to employ are based on the anticipated effort involved in collecting data and other case-specific criteria (De Jesus

*Corresponding author: boydnbarrett@gmail.com
Received September 25, 2016; accepted April 24, 2017

et al. 2011). Once a survey method is chosen, rigorous sample size planning should be undertaken to ensure that estimates are sufficiently accurate and precise to allow for detection of changes over time (Kelley and Rausch 2006). Sample size planning involves a trade-off analysis wherein a researcher concurrently assesses multiple factors, such as survey administration cost, accuracy and precision of estimates, the magnitude and direction of biases, and variance in the statistic of interest (Lenth 2001).

Off-site fishery surveys are a practical alternative to on-site surveys when the angler population is dispersed spatially and temporally, as is the case for many recreational fisheries (Roach et al. 1999; Hartill et al. 2012). However, off-site angler surveys are prone to a low response rate and to reporting bias (Roach et al. 1999; Zarauz et al. 2015). Surveys with a low response rate are at risk of producing biased population estimates (Brick 2013) and should be viewed with skepticism if a nonresponse bias assessment is not undertaken (Fisher 1996; Lewis et al. 2013). In essence, the potential benefits of off-site angler surveys, such as cheaper cost per contact (McCormick et al. 2015), can easily be negated by nonresponse bias. It is therefore imperative to explore the effects of nonresponse when making a decision to use off-site angler surveys (McCormick et al. 2015).

Nonresponse bias assessments involve the use of additional survey effort to contact sample members who have a low propensity to respond (Peress 2010) and statistical tests to determine whether there is a significant difference between respondents and nonrespondents with regard to the variables of interest (Clottey and Benton 2013). Nonresponse bias in angler surveys occurs when respondents are more avid anglers who have more successful fishing trips or expend more effort on the fishery than nonrespondents (Connelly et al. 2000; Zarauz et al. 2015). Simply extrapolating the mean without any correction from a data set in which avid anglers are disproportionately represented will result in an overestimation of population parameters (Fisher 1996; Connelly et al. 2000). It is important to note that nonresponse bias does not always exist in surveys with a low response rate (Larkin et al. 2010). However, researchers can only make a conclusion about the presence or absence of nonresponse bias after undertaking a nonresponse bias assessment. Unfortunately, nonresponse bias assessments usually have a higher cost per contact compared to initial forms of contact (Peytchev 2013) and can easily put surveys over budget or result in situations where a less-than-ideal sample size is used.

Survey mode selection is an important component of the survey design process (De Leeuw 2005). Obvious considerations pertaining to survey mode include its effect on administration cost and how to harness the distinctive advantages provided by different modes (e.g., fast turnaround time for e-mail surveys; De Leeuw 2005; Buelens and van den Brakel 2015). However, a less-obvious and often-overlooked consideration is the fact that survey mode can potentially prompt varied responses from survey participants and introduce a mode-effect bias (Vannieuwenhuyze and Loosveldt 2013; Wallen et al. 2016). Mode-effect is a type of measurement error (Vannieuwenhuyze et al. 2010; Dillman et al. 2014) and has been shown to have a serious effect on population estimates (Laborde et al. 2014; Zarauz et al. 2015). For example, a mail survey, e-mail survey, and telephone survey of anglers targeting European Sea Bass *Morone labrax* yielded total annual catch estimates of 351, 129, and 156 metric tons of fish, respectively (Zarauz et al. 2015). The attractiveness of using a single survey mode to disseminate a questionnaire might prove to be misleading if that mode collects data disproportionally from the sample or evokes a particular response pattern, thus calling for mixed-mode designs (Wallen et al. 2016).

The two primary statistical analyses in sample size planning are power analysis and accuracy in parameter estimation (AIPE; Kelley and Rausch 2006). Power analysis aims to determine the sample size necessary to detect a particular effect size given a chosen type I error rate (Peterman 1990). Power analyses are undertaken when the aim of a study is to conduct hypothesis testing. Alternatively, AIPE, which is rarely used in fisheries research, determines the sample size necessary to produce a predefined precision in a population parameter (Kelley et al. 2003; Kelley 2007). The decision on which analysis to use is reliant on the aim of the study (Kelley et al. 2003). As fisheries studies are more commonly concerned with model precision and its effect on reference points and management decisions than testing a particular hypothesis, AIPE is arguably more appropriate in a fisheries context.

Effort and catch estimates have an intuitive meaning and can be represented by a point estimate and a confidence interval (Vaske 2002). Surveys that yield estimates with high uncertainty and low accuracy have little scientific value and can lead to dangerously misleading management advice (Kelley et al. 2003; Kelley 2007). Rigorous and systematic sample size planning is therefore needed to determine the minimum sample size that will obtain the desired level of precision (Lenth 2001) and the most suitable survey design alternative that will maximize accuracy—all while keeping the survey within budget (De Leeuw 2005). A simultaneous assessment of how sample size and survey design affect accuracy and precision can be undertaken in the form of a sensitivity analysis, where changes in dependent variables are assessed based on changes in independent variables.

We used effort and catch data from an annual off-site survey administered to anglers targeting White Sturgeon *Acipenser transmontanus* to illustrate how a sensitivity analysis can be used for sample size planning. The actual survey was administered via three modes: mail, e-mail, and telephone. We used nearly identical questionnaires for each mode (i.e., the only differentiating factor was that the e-mail questionnaire had six additional demographic questions), permitting a comparison of the relative bias and cost effectiveness of various combinations of the three modes. We had three specific aims. Our first aim was to determine the effect of nonresponse bias on total annual effort and catch estimates. Second, we investigated the effect of survey mode on sample estimates. Third, we used survey effort and catch data in a sensitivity analysis to determine the variation in total annual effort and catch estimates as a function of sample size and survey mode. The sensitivity analysis mimicked the actual data collection procedure. Although our results are specific to the White Sturgeon case study, the methods used are instructive for anyone looking to evaluate and improve the administration of off-site surveys through appropriate sample size planning to inform fisheries management.

METHODS

*Study system.—*There are six recognized populations of White Sturgeon in British Columbia. Four populations are legally listed under Schedule 1 of the Species at Risk Act. The Committee on the Status of Endangered Wildlife in Canada (COSEWIC) assessed the other two populations as endangered (COSEWIC 2012). The two endangered populations (i.e., lower and middle Fraser River populations; Figure 1) are targeted in a popular catch-and-release recreational fishery and are the focus of this study.

Anglers who participate in the lower Fraser River White Sturgeon recreational fishery (hereafter, “sturgeon anglers”) are legally obligated to purchase a White Sturgeon Conservation License (WSCL) in addition to a basic provincial angling license. Sturgeon anglers originate from within British Columbia, across Canada, and internationally. The WSCL is valid for one fishing season (between April 1 and March 31 of the following year). Holders of a WSCL either fish with the assistance of a licensed White Sturgeon guide (guided anglers) or without a licensed guide (nonguided anglers).

The Annual Fraser River White Sturgeon Angling Questionnaire (hereafter, “sturgeon angler survey”) is used to collect data from WSCL holders (see Figure 2 for a copy of the questionnaire). The British Columbia Ministry of Forests, Lands, and Natural Resource Operations (FLNRO) administers the sturgeon angler survey at the end of each fishing season. The survey asks WSCL holders to report their angling status (i.e., guided angler, nonguided angler, or did not fish) for the license year; licensed guides report guided anglers’ effort and catch data via separate mandatory reporting. Subsequently, only nonguided anglers are asked to report their annual effort and annual catch within each management region. The survey aims to estimate the total annual catch, total annual effort, and annual CPUE for nonguided sturgeon anglers. Within the context of the sturgeon angler survey, effort is expressed in angler-days, and catch is the number of White Sturgeon that are caught and released by an individual angler. We defined an angler-day as an entire day or any part of a day spent fishing for White Sturgeon.

*Data collection.—*This research used data from the 2014–2015 sturgeon angler survey. Prior to 2014–2015, catch and effort data were assessed annually via mail surveys that were sent to all WSCL holders. We made changes to the original sturgeon angler survey to meet the requirements of our statistical tests and simulation exercise. First, we selected a simple random sample from the sturgeon angler population. Second, we added an e-mail survey and a telephone survey to the survey methodology. Third, we selected a sample to act as a control group for the e-mail survey (as explained below).

The 2014–2015 sturgeon angler survey was a mixed-mode survey with three modes: mail, e-mail, and telephone (see Figure 2 for a copy of the questionnaire). We used the tailored design method (TDM) to guide survey design (Dillman et al. 2014). The TDM recommends using the same question across all survey modes and using similar visual formats, among other design strategies geared toward reducing measurement differences across survey modes (Dillman et al. 2014). The e-mail survey had some additional demographic questions at the end of the questionnaire. These additional questions were not mandatory, and the responses of anglers that chose to discontinue answering questions at that point were still included in our analysis. The survey followed a two-phase sample design (Lukacs 2007). The sample unit was an individual angler identified by their license number. In the first phase, we selected a random sample from the population of anglers who purchased a WSCL for the 2014–2015 fishing season. In the second phase, we subsampled from the group of nonrespondents of the first phase. We intentionally selected large samples. We used the following procedure to select three independent samples from the sturgeon angler population. First, we divided the angler population (N = 15,471) into two mutually exclusive groups: the L_{con} group (n = 6,549) consisted of anglers who provided an e-mail address and consented to being contacted via e-mail when purchasing their WSCL, and the $L_{no.con}$ group (n = 8,922) consisted of all other anglers. Second, two simple random samples were concurrently selected from the L_{con} angler group: anglers belonging to $L_{con(e)}$ (n = 4,248) each received a questionnaire via e-mail, and those belonging to $L_{con(m)}$ (n = 1,098) each received a questionnaire via mail. Third, we selected a simple random sample from the $L_{no.con}$ group, who received a survey by mail. We used $L_{con(m)}$ as a control group for the e-mail survey ($L_{con[e]}$) to compare response rate and sample representativeness related to survey mode. Each angler was eligible to receive only one questionnaire regardless of the number of licenses purchased. For example, an angler who purchased multiple 1-d licenses would only be sent one questionnaire. We asked participants to report their annual effort and annual catch for all licenses on the questionnaires they received to reduce the likelihood of double reporting. In essence, we had

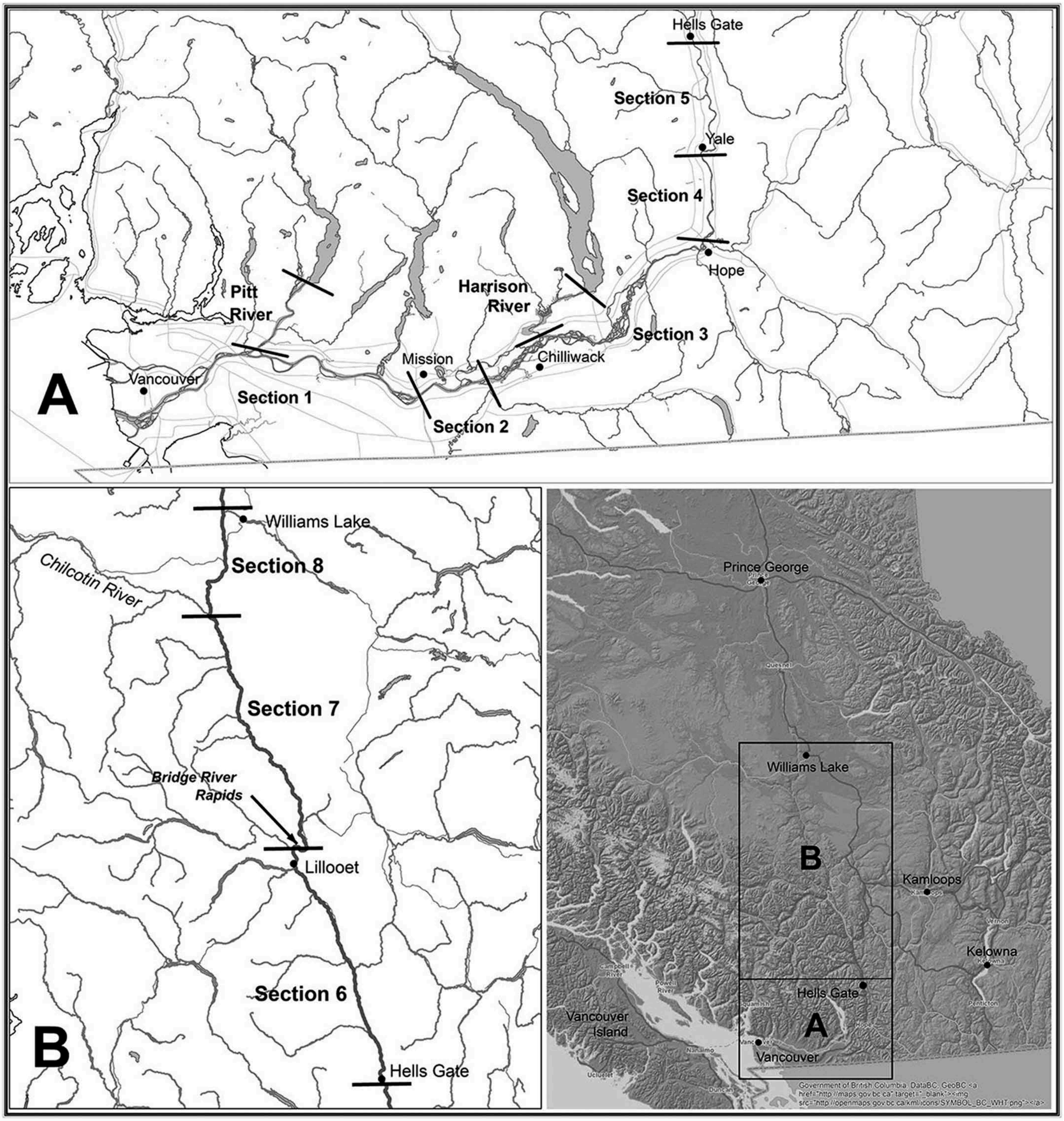
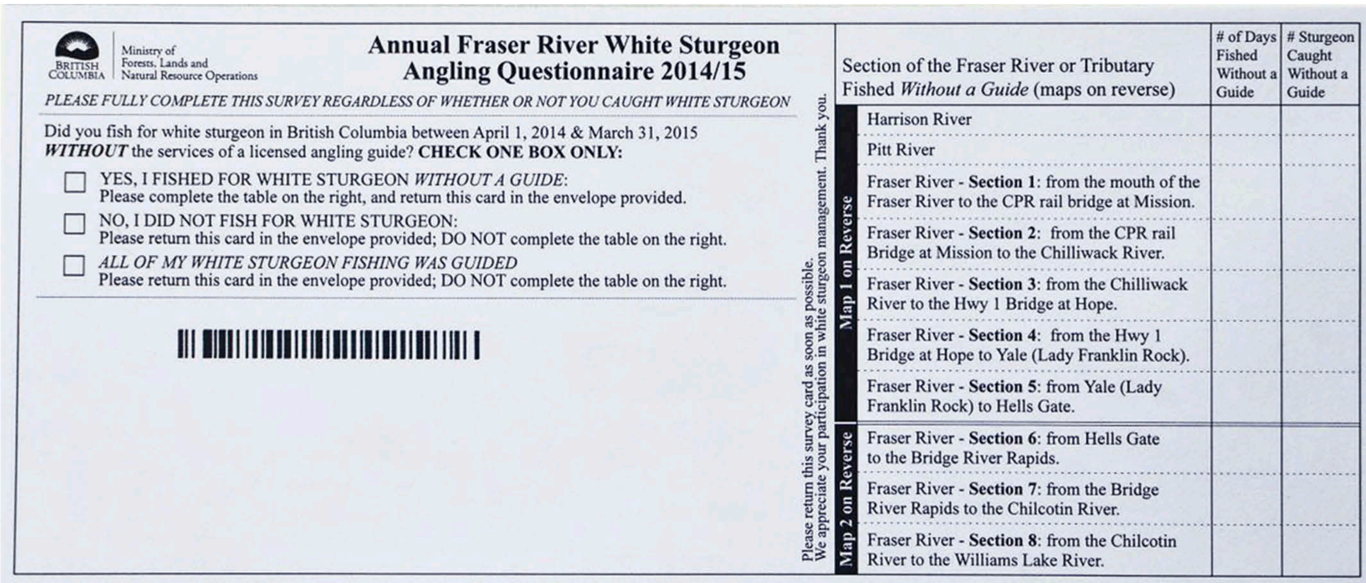


FIGURE 1. Image of the study area (Fraser River, British Columbia), where surveyed anglers targeted White Sturgeon (section numbers correspond to those listed in the survey questionnaire; Figure 2). Boxes (A) and (B) indicate the general location of the study area. The image is not drawn to scale.

three survey groups ($L_{con[e]}$, $L_{con[m]}$, and $L_{no.con}$) and three survey modes (mail, e-mail, and telephone).

The mail survey (n = 4,470) was sent to anglers from $L_{con(m)}$ (n = 1,098) and $L_{no.con}$ (n = 3,372). We sent the first invitation to participate on June 15, 2015, and a reminder notification 6 weeks later. Each Canadian participant package included the paper questionnaire and a stamped return envelope in both the initial and reminder letters. Anglers



BRITISH COLUMBIA | Ministry of Forests, Lands and Natural Resource Operations

Annual Fraser River White Sturgeon Angling Questionnaire 2014/15

PLEASE FULLY COMPLETE THIS SURVEY REGARDLESS OF WHETHER OR NOT YOU CAUGHT WHITE STURGEON

Did you fish for white sturgeon in British Columbia between April 1, 2014 & March 31, 2015 ***WITHOUT*** the services of a licensed angling guide? **CHECK ONE BOX ONLY:**

☐ YES, I FISHED FOR WHITE STURGEON *WITHOUT A GUIDE*: Please complete the table on the right, and return this card in the envelope provided.

☐ NO, I DID NOT FISH FOR WHITE STURGEON: Please return this card in the envelope provided; DO NOT complete the table on the right.

☐ *ALL OF MY WHITE STURGEON FISHING WAS GUIDED* Please return this card in the envelope provided; DO NOT complete the table on the right.

Please return this survey card as soon as possible. We appreciate your participation in white sturgeon management. Thank you.

	Section of the Fraser River or Tributary Fished *Without a Guide* (maps on reverse)	# of Days Fished Without a Guide	# Sturgeon Caught Without a Guide
	Harrison River		
	Pitt River		
Map 1 on Reverse	Fraser River - **Section 1**: from the mouth of the Fraser River to the CPR rail bridge at Mission.		
Map 1 on Reverse	Fraser River - **Section 2**: from the CPR rail Bridge at Mission to the Chilliwack River.		
Map 1 on Reverse	Fraser River - **Section 3**: from the Chilliwack River to the Hwy 1 Bridge at Hope.		
Map 1 on Reverse	Fraser River - **Section 4**: from the Hwy 1 Bridge at Hope to Yale (Lady Franklin Rock).		
Map 1 on Reverse	Fraser River - **Section 5**: from Yale (Lady Franklin Rock) to Hells Gate.		
Map 2 on Reverse	Fraser River - **Section 6**: from Hells Gate to the Bridge River Rapids.		
Map 2 on Reverse	Fraser River - **Section 7**: from the Bridge River Rapids to the Chilcotin River.		
Map 2 on Reverse	Fraser River - **Section 8**: from the Chilcotin River to the Williams Lake River.		

FIGURE 2. A copy of the Annual Fraser River White Sturgeon Angling Questionnaire used to collect data from anglers targeting White Sturgeon in British Columbia. The questionnaire was used to collect data for the mail survey (see Figure 1 for section locations). We used similar questions for the e-mail and telephone surveys.

with an address outside of Canada paid their own return postage. Envelopes were delivered by standard mail.

The e-mail survey was sent to anglers from $L_{con(e)}$ ($n = 4{,}248$). We sent the first invitation to participate on June 23, 2015, and a reminder e-mail 14 d later. Each e-mail was personalized and had a hyperlink that allowed anglers to connect directly to the survey; responses were automatically stored to a database.

The nonresponse assessment survey (hereafter, "follow-up survey") was administered via telephone during the first week of December 2015. Only anglers who provided a Canadian address when purchasing their WSCL were contacted. We selected a simple random sample (n = 1,000) from the list of anglers who were originally sent one of the three surveys but did not respond to $L_{con(e)}$ (n = 582), $L_{no.con}$ (n = 314), and $L_{con(m)}$ (n = 104). Four attempts were made to contact an angler before that angler was deemed a "no contact." Calls were made between 1700 and 2100 hours on weekdays and between 1000 and 2000 hours on weekends.

Analyses.—We calculated the adjusted response rate (R) for each survey mode as

$$R = \frac{n_c}{n_s - U} \times 100\%, \quad (1)$$

where n_c is the number of completed questionnaires returned, n_s is the total sample (number of questionnaires sent), and U is the number of undeliverable questionnaires. We defined a completed return as a questionnaire that had a single, legible response for each required question (except for the demographic questions added to the e-mail survey). We estimated the percentage of noncontact in the e-mail survey and the mail survey by asking participants in the follow-up survey if they recalled receiving a copy of the White Sturgeon questionnaire by e-mail or mail during the first contact. We defined noncontacts as sample members who stated that they did not receive a survey via e-mail or mail during the first phase of contact.

We used the R statistical language to analyze the data (R Development Core Team 2014). We tested for nonresponse bias within each survey group by using a chi-square test to compare the proportion of anglers (i.e., guided anglers, nonguided anglers, and anglers who did not fish) in the first contact with proportions in the follow-up survey. Additionally, we used Welch's tests on rank-transformed data (Ruxton 2006; Zimmerman 2012) within each survey group to test for a difference between annual effort and annual catch estimates obtained from nonguided anglers in the first contact with estimates from the same group in the follow-up survey. We used a third series of Welch's tests on rank-transformed data to test for a difference between total effort and total catch estimates obtained from anglers in the three survey modes (mail, e-mail, and telephone). We used Dunn's multiple-comparison post hoc tests for pairwise comparisons when the Welch's test provided evidence of a difference. The type I error rate (α) was set at 0.05 for all main-effect tests, and a Bonferroni adjustment was used for post hoc analyses.

Simulation.—Our simulation exercise was in the form of a sensitivity analysis. We performed the sensitivity analysis using

R (Wickham 2009; R Development Core Team 2014; Baptiste 2016; Wickham and Francois 2016) to determine the precision and accuracy of total annual effort and catch estimates and the variation in survey cost as a function of sample size and four different survey design alternatives. The survey design alternatives varied based on the number of survey modes and the presence or absence of data from follow-up surveys. Survey alternatives were as follows: (1) all survey modes (data from all survey modes were used); (2) first contact only (only those data from the first contacts [mail and e-mail] were used); (3) e-mail survey only (only those data from the first contact of the e-mail survey were used); and (4) mail survey only (only those data from the first contact of the mail survey were used). The angler survey data set included completed questionnaires, incomplete questionnaires, and no contacts. We undertook the simulation exercise by resampling without replacement from the 2014–2015 angler survey data set. We repeated each survey alternative at nine different sample sizes that ranged from 20% to 100% of the original sample size at intervals of 10%. Each sample size was resampled over 1,000 iterations. Estimates of total annual effort ($\hat{E}$) and total annual catch ($\hat{C}$) were calculated for each new sample as

$$\hat{E} = p_{Ng} \times N \times \bar{E}, \quad (2a)$$

$$\hat{C} = p_{Ng} \times N \times \bar{C}, \quad (2b)$$

where p_{Ng} is the proportion of nonguided anglers in the responding sample; N is the total number of White Sturgeon anglers (known from license sales); $\bar{C}$ is the mean catch per nonguided angler; and $\bar{E}$ is the mean effort per nonguided angler.

The cost of administering the survey was calculated for each design alternative. The costs of the mail survey and the telephone survey were broken down to a per-angler basis. The cost per angler included the cost per contact and the cost of data entry (in Canadian dollars). The e-mail survey had a fixed cost regardless of the sample size. We multiplied cost per contact by the sample size to estimate total cost for each survey design alternative.

We evaluated the performance of each survey design alternative by assessing the associated bias, precision, accuracy, and survey cost against the estimates obtained from the original sturgeon angler survey using all survey modes. We used the coefficient of variation (CV) to assess precision and the root mean square error to evaluate accuracy (following Walther and Moore 2005; Table 1).

RESULTS

Among the three survey modes, the adjusted response rate was lowest for the e-mail survey (22%; Table 2). Comparatively, response rates for the two mail surveys and three telephone surveys fell between 39% and 44%. Our telephone follow-up survey indicated that the e-mail survey had a higher percentage of noncontact (63%) than the two mail surveys ($L_{no.con}$: 44%; $L_{con[m]}$: 40%).

Our first test for nonresponse bias utilized a chi-square test to determine whether the proportions of anglers (i.e., guided, nonguided, and anglers who did not fish) among first-contact respondents were equal to the proportions among follow-up contact respondents within the three survey groups: $L_{con(e)}$, $L_{con(m)}$, and $L_{no.con}$. The groups $L_{con(e)}$ ($\chi^2 = 29.6$, df = 2, $P < 0.01$) and $L_{no.con}$ ($\chi^2 = 7.2$, df = 2, $P = 0.03$) were both found to have significantly different proportions. In each case, the proportion of anglers who responded that they did not fish was higher in the respective follow-up survey than in the first contact, suggesting that anglers who did not catch White Sturgeon were less compelled to complete the survey using these modes. In contrast, the percentage of $L_{con(m)}$ anglers who reported that they did not fish was greater at the first contact than in the follow-up survey, but the result from the chi-square test was not significant.

There was high inequality in catch and effort among anglers. Twenty-five percent of nonguided anglers accounted for 85% of the total annual effort and 72% of the total annual catch reported for all survey modes. The medians of total effort and total catch reported by individual nonguided anglers

TABLE 1. Notation and equations for coefficient of variation (CV) and root mean square error (RMSE) that were used to assess bias, precision, and accuracy, respectively, in the survey of White Sturgeon anglers. Equations were adapted from Walther and Moore (2005).

Equation	Description	Equation number
$\mathrm{CV} = \frac{\mathrm{SD}}{\bar{E}}$	The CV was used to assess the precision associated with each survey design alternative (SD is the standard deviation; $\bar{E}$ is the mean of estimates E across 1,000 iterations).	3a
$\mathrm{RMSE} = \sqrt{\frac{1}{n} + \sum_{j=1}^{n}(E_j - A)^2}$	The RMSE was used to assess the accuracy associated with each survey design alternative (n is the number of samples; Σ is the summation formula; E is the estimate [effort or catch] obtained from the jth sample; A is the annual total effort or annual total catch obtained from the actual sturgeon angler survey).	3b

TABLE 2. Table showing adjusted response rates for the three White Sturgeon angler survey groups: $L_{con(e)}$ (the e-mail group); $L_{con(m)}$ (the mail group with anglers who provided an email address); and $L_{no.con}$ (the mail group with anglers who did not provide an email address). The e-mail survey had the lowest adjusted response rate (22%). Comparatively, response rates for the mail and telephone surveys ranged from 39% to 44%. Data are results from the 2014–2015 Annual White Sturgeon Angling Questionnaire survey. We used the Hmisc package in R (Harrell et al. 2016) to calculate the binomial 95% confidence limits (CLs) of the response rate.

Survey group and mode	Contact phase	Sample size	Response rate (%)	Lower 95% CL	Upper 95% CL
$L_{con(e)}$					
E-mail	First contact	4,248	22	20	23
Telephone	Follow-up contact	582	43	39	47
$L_{con(m)}$					
Mail	First contact	1,098	44	41	47
Telephone	Follow-up contact	104	41	32	51
$L_{no.con}$					
Mail	First contact	3,372	41	39	43
Telephone	Follow-up contact	314	39	34	44

were 3 angler-days (range = 1–120 angler-days) and 3 White Sturgeon (range = 0–500 fish), respectively.

We observed significant differences in effort estimates (Welch's test: $F_{2,\ 227.4} = 5.64$, $P < 0.01$) and catch estimates (Welch's test: $F_{2,\ 224.6} = 9.52$, $P < 0.01$) from nonguided anglers as a function of survey mode. Post hoc analyses indicated that this difference occurred between the e-mail and mail surveys for individual annual effort ($P < 0.01$; e-mail: median = 4 angler-days; mail: median = 3 angler-days) and individual annual catch of White Sturgeon ($P < 0.01$; e-mail: median = 4 fish; mail: median = 2 fish). Post hoc tests did not provide evidence of a clear difference between the telephone survey and either the e-mail survey or the mail survey.

Our second method of evaluating nonresponse bias utilized a Welch's test to compare effort and catch estimates obtained from first-contact respondents with those from follow-up-contact respondents within the survey groups $L_{con(e)}$, $L_{con(m)}$, and $L_{no.con}$. Only one of the six Welch's tests, the test of $L_{con(m)}$ catch, yielded statistically significant results (Table 3). Results from the Welch's test indicated that $L_{con(m)}$ follow-up contact respondents tended to catch more White Sturgeon than $L_{con(m)}$ first-contact respondents (follow-up: n = 13, mean rank = 64.5; first contact: n = 79, mean rank = 43.5).

Sample Size Planning

Data collected during the White Sturgeon survey were resampled to demonstrate how accuracy, precision, bias, and survey cost varied as a function of sample size and various combinations of survey modes (mail, e-mail, and telephone). When the all-survey-modes alternative was used, accuracy of total annual effort and total annual catch declined with sample size, but estimates were unbiased (Figures 3, 4, left panels). There was a minor loss in accuracy of catch estimates—and an even smaller loss in accuracy of effort estimates—at lower sample sizes. Similarly, uncertainty in estimates of effort and catch increased as sample size declined, as evidenced by the CVs. Survey cost linearly decreased with sample size but was always higher than the cost of the other alternatives at the same sample size.

The first-contact-only alternative (i.e., including mail and e-mail surveys) yielded effort estimates that were similar to those from the all-survey-modes alternative in terms of accuracy, bias, and precision; survey costs were lower (Figure 3, second column of panels). However, total annual catch estimates were negatively biased at all sample sizes.

TABLE 3. Results of Welch's tests comparing ranked estimates from the first contact and follow-up survey for each White Sturgeon angler survey group (group symbols are defined in Table 2). Data were obtained from the 2014–2015 Annual Fraser River White Sturgeon Angling Questionnaire survey. The survey followed a two-phase sample design; data from the two phases (i.e., within each survey group) were compared to test for nonresponse bias. Type-I error rate (α) was set at 0.05. Only catch estimates for $L_{con(m)}$ were significantly different.

	Statistic		
Survey group and variable	df	F-value	P-value
$L_{con(e)}$			
Effort	1, 58.47	2.89	0.09
Catch	1, 59.10	0.27	0.60
$L_{con(m)}$			
Effort	1, 14.88	2.68	0.12
Catch	1, 17.34	8.97	0.01
$L_{no.con}$			
Effort	1, 32.57	0.71	0.41
Catch	1, 31.99	0.05	0.83

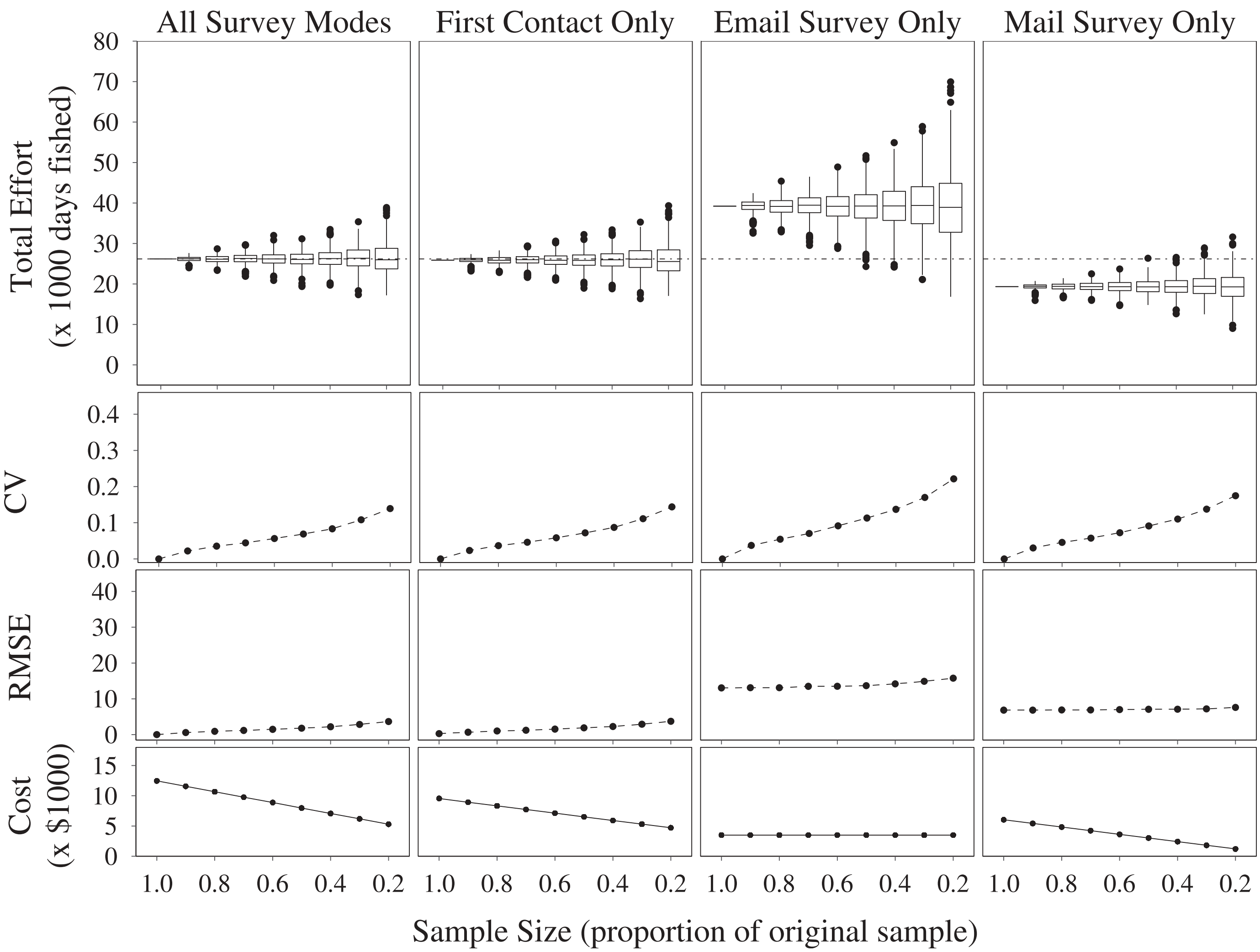


FIGURE 3. Point estimates and distributions of estimates for total annual effort, coefficient of variation (CV), root mean square error (RMSE), and survey cost (CAN$) as a function of angler sample size based on the simulation model. Data (for nonguided anglers only) were randomly selected without replacement from the 2014–2015 Annual Fraser River White Sturgeon Angling Survey data set. Results of simulations using data from all surveys combined, e-mail survey only, first contact only, and mail survey only are presented. The horizontal dashed line in the upper panels represents the true point estimate of effort based on the original complete survey.

The e-mail-survey-only alternative had a low and consistent survey cost but yielded the largest CV and RMSE (Figures 3, 4, third column of panels). The precision of estimates for this alternative decreased with sample size faster than that for the other survey mode alternatives. Both total annual effort and total annual catch were positively biased at all sample sizes and resulted in consistently higher effort and catch estimates than the other survey mode alternatives.

The mail-survey-only alternative had the lowest total survey cost at smaller sample sizes but consistently underestimated total annual effort and total annual catch at all sample sizes. Loss of precision for the mail-survey-only alternative as sample size declined was comparable to that observed for the all-survey-modes and first-contact-only alternatives, as was the decrease in precision (i.e., CV) for estimates.

DISCUSSION

We used four criteria (i.e., bias, precision, accuracy, and cost) to evaluate single- and mixed-mode survey designs in an off-site angler survey. We evaluated an e-mail-only survey, a mail-only survey, and various combinations of e-mail, mail, and telephone surveys. We found that (1) response rates were substantially lower for the e-mail survey than for mail and telephone surveys; (2) nonresponse bias had an effect on total annual effort and total annual catch estimates, but the effect

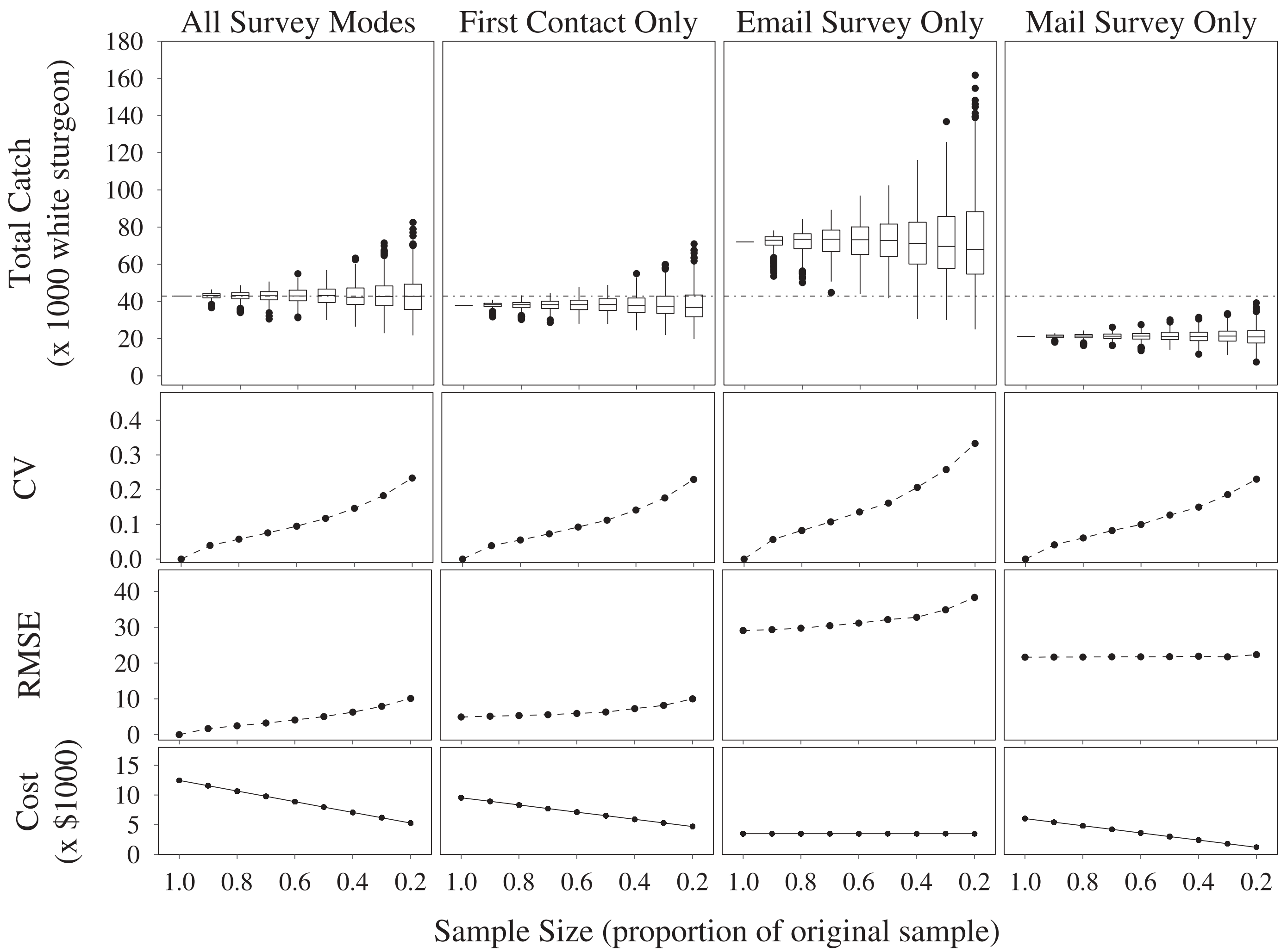


FIGURE 4. Point estimates and distributions of estimates for total annual White Sturgeon catch, coefficient of variation (CV), root mean square error (RMSE), and survey cost (CAN$) as a function of angler sample size based on the simulation model. Data (for nonguided anglers only) were randomly selected without replacement from the 2014–2015 Annual Fraser River White Sturgeon Angling Survey data set. Results of simulations using data from all surveys combined, e-mail survey only, first contact only, and mail survey only are presented. The horizontal dashed line in the upper panels represents the true point estimate of catch based on the original survey.

was small and was inconsistent across the three survey groups; and (3) the resulting estimates (total effort and total catch) from the mail and e-mail survey modes were significantly different. The high variability of total catch between anglers compounded the effect that sample size reduction had on precision.

A limitation of this research arose from the fact that we did not obtain a 100% response rate in the follow-up surveys (i.e., the telephone surveys). During the survey planning phase, we assumed that the random sample taken from the group of anglers who did not respond to the first contact would be representative of all anglers who did not respond. With only a 44% response rate in the follow-up survey, there was a group of anglers for whom we did not obtain any data. These follow-up survey nonrespondents might have had different effort and catch patterns from the follow-up survey respondents. Therefore, our follow-up survey provided insight into—as opposed to an irrefutable conclusion about—the characteristics of the group of nonrespondents to the first phase of contacts (e-mail and mail surveys) and the general direction and magnitude of nonresponse bias. Another limitation arose from the fact that the e-mail survey had more questions (six additional demographic questions) than the mail survey. The fact that the e-mail survey was longer could have caused the low response rate. However, we do not expect this effect to be substantial because the additional questions were at the end of the survey and were optional. Respondents were able to submit their annual effort and annual catch data without completing the demographic questions.

Our e-mail survey had the fastest turnaround time and lowest cost per contact but yielded a lower response rate and had an incomplete sample frame, as has been reported elsewhere (e.g., Gigliotti and Dietsch 2014; Zarauz et al. 2015; Wallen et al. 2016). The cost advantage of the e-mail survey makes it an attractive alternative (McCormick et al. 2015); therefore, it is important to explore the reasons for its poor performance before dismissing it as an unsuitable mode (Gigliotti and Dietsch 2014). Potential causes of the poor e-mail survey performance include the presence of invalid, incorrect, or old e-mail addresses in the data frame; technical problems with the communication path of emails; and personal and organizational e-mail settings that may have identified survey emails as spam (Callegaro et al. 2015). A large (63%) proportion of anglers who did not respond to our e-mail survey noted in the telephone follow-up survey that they did not receive an e-mail invitation. This finding corresponds with that of Callegaro et al. (2015) and suggests a need to double-check spelling or verify e-mail validity during license purchase. The response rate of the e-mail survey would have been markedly higher if we had correct contact information. Incorporating e-mail verification into the license purchase process will improve the quality of the database over the long term and allow survey designers to take full advantage of e-mail surveys.

Nonresponse bias occurs when there is a systematic difference between respondents and nonrespondents with regard to key survey variable(s) of interest (Connelly et al. 2000). We did not observe a consistent difference between respondents and nonrespondents regarding estimates of means (effort and catch) or proportions (guided anglers, nonguided anglers, and anglers who did not fish). Only one of six survey mode combinations for effort and catch showed a difference for catch between respondents and nonrespondents. The other potential cause of nonresponse bias that we observed (i.e., anglers who did not fish were less likely to respond) did not have a meaningful effect on total annual catch and total annual effort estimates since we did not use those values in our calculations. We only used estimates reported by nonguided anglers. One common cause of nonresponse bias is an overrepresentation of avid anglers in the responding sample (Fisher 1996; Connelly et al. 2000; Zarauz et al. 2015). However, we did not observe this overrepresentation of avid anglers in the group of nonguided angers who responded. We theorize that nonresponse bias was minimal because nonguided sturgeon anglers are a specialized angler group. Nonguided anglers who participate in the White Sturgeon catch-and-release fishery need specialized equipment and a unique license, and they use a somewhat distinctive method when fishing to catch a single species; these are characteristics of a specialized fishery or specialized anglers (Fisher 1997; Landsman et al. 2011). Specialized anglers are typically vocal about management issues related to their target species and are particularly concerned with the quality of the fishery (Beardmore et al. 2014). It can be argued that majority of nonguided sturgeon anglers in our survey were specialized anglers who had a keen interest in the fishery's quality and were willing to participate in our survey regardless of their catch or participation rate. We surmise that there is an increased likelihood for the occurrence of nonresponse bias as angler specialization decreases across a fishery. Managers of fisheries with more diverse levels of angler specialization (e.g., Hutt and Bettoli 2007) should explore survey results for nonresponse bias since the results observed here might not hold true in fisheries with less-specialized anglers.

Survey mode can have a profound effect on the data reported by participants (Vannieuwenhuyze and Loosveldt 2013; Laborde et al. 2014; Wallen et al. 2016). The benefit of a mixed-mode survey design is that each mode tends to negate the bias produced by other modes. Transitioning to an e-mail-only or mail-only survey design will lead to incorrect estimation of both total annual effort and total annual catch, possibly prompting a management response when none is warranted or vice versa. However, e-mail survey designs might still have a place in off-site angler surveys if the aim is to collect data from a large sample (Zarauz et al. 2015). Additionally, consider a situation in which a researcher's key interest lies in assessing relative effort and catch trends as opposed to absolute values. In this situation, the researcher can use an e-mail-only survey design with the assumption that the direction and magnitude of the bias remain constant for each survey. Furthermore, this assumption could be verified using follow-up surveys. Understanding the relative accuracy and precision of effort and catch estimates from e-mail surveys a priori will be important in this decision.

Our simulation exercise went beyond AIPE analyses because AIPE evaluates parameter uncertainty against sample size (Kelley 2007), while our analysis also examined the underlying factors that affect sample size (e.g., methods to control for nonresponse bias and associated survey cost) and the effect of those factors on accuracy and precision of parameters. These measures of survey performance are important to consider when evaluating survey design (De Leeuw 2005). For example, we found that the change in accuracy as a function of survey cost suggests that small sample sizes could be permitted for a mail-only survey. However, when examining other metrics for the mail-only survey, we found that catch precision was quite low, and both effort and catch were negatively biased. With these results, it may be difficult to justify mail-only surveys. Our work highlights the need to incorporate multiple metrics in the sample size planning process, as this will help when assessing alternatives.

Uncertainty around an estimate is as important as the estimate itself (Kelley 2007). The AIPE approach urges survey designers to plan surveys with the aim of obtaining estimates with an acceptable level of uncertainty based on the needs of the study (Kelley et al. 2003; Kelley 2007). Our simulation exercise indicated that we could not undertake any meaningful reduction in sample size (i.e., a reduction that resulted in substantial cost savings) without substantially increasing variance. This suggests the need for a more sophisticated experimental design, such as a stratified random sample. Given the high variability of sturgeon anglers' effort and catch estimates, a shift from a simple random sample to a stratified random sample design is apparently needed. Stratification can reduce the variance of estimates (Holt and Smith 1979). However, a successful stratified design would need to ensure that anglers are stratified based on factors that (1) affect their effort and catch and (2) minimize variability within each stratum. We did not have enough data to test the effect of a stratified random sample. However, we theorize that sturgeon anglers can be stratified based on license type (1 d, 8 d, or annual) or number of years of fishing for White Sturgeon since this information is currently known from the license purchasing history. Stratified random designs should be included in future sample size planning.

Our research brings into focus the potential cost savings that can be achieved with minimal influence on effort and catch estimates if rigorous sample size planning is undertaken during survey planning. A researcher, fishery manager, or fish biologist can incorporate a similar simulation exercise into a decision-making framework and explore the sensitivity of population parameters to factors such as survey cost, sample size, magnitude and direction of biases, and the number and types of survey modes. However, it will be necessary to decide on an acceptable level of error so that a sample size can be determined. We also demonstrate that survey mode can affect sample estimates and that using a single, cheap survey mode is not always ideal if the aim is to identify total effort and catch. Our research shows that there is a need for sample size planning even in a relatively small, gear-specific fishery with a complete sample frame. The need for sample size planning is therefore even more important in complex fisheries, especially when a low level of error is desired and the survey must be undertaken within a tight budget.

ACKNOWLEDGMENTS

The British Columbia FLNRO provided data for this project. The project was initiated and assisted throughout by Duane Jesson (FLNRO), to whom we are indebted. We are grateful to the Habitat Conservation Trust Foundation, and the anglers, hunters, trappers, and guides who contribute to the trust, for making a significant financial contribution to support this research project. Without such support, this project would not have been possible.

REFERENCES

Baptiste, A. 2016. gridExtra: miscellaneous functions for "Grid" graphics. R package version 2.2.1. R Foundation for Statistical Computing, Vienna. Available: http://cran.r-project.org/package=gridExtra. (May 2017).

Beardmore, B., L. M. Hunt, W. Haider, M. Dorow, and R. Arlinghaus. 2014. Effectively managing angler satisfaction in recreational fisheries requires understanding the fish species and the anglers. Canadian Journal of Fisheries and Aquatic Sciences 72:500–513.

Brick, J. M. 2013. Unit nonresponse and weighting adjustments: a critical review. Journal of Official Statistics 29:329–353.

Buelens, B., and J. A. van den Brakel. 2015. Measurement error calibration in mixed-mode sample surveys. Sociological Methods and Research 44:391–426.

Callegaro, M., K. L. Manfreda, and V. Vehovar. 2015. Web survey methodology. SAGE Publications, London.

Clottey, T., and W. C. Benton. 2013. Guidelines for improving the power values of statistical tests for nonresponse bias assessment in OM research. Decision Sciences 44:797–812.

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 5:19–29.

COSEWIC (Committee on the Status of Endangered Wildlife in Canada). 2012. COSEWIC assessment and status report on the White Sturgeon *Acipenser transmontanus*in Canada. COSEWIC, Gatineau, Quebec.

De Jesus, M. J., S. J. Magnelia, and C. C. Bonds. 2011. Use of a volunteer angler survey for assessing length distribution and seasonal catch trends of trophy Largemouth Bass. North American Journal of Fisheries Management 29:1225–1231.

De Leeuw, D. 2005. To mix or not to mix data collection modes in surveys. Journal of Official Statistics 21:233–255.

Dillman, D. A., J. D. Smyth, and L. M. Christian. 2014. Internet, phone, mail, and mixed-mode surveys: the tailored design method. Wiley, Hoboken, New Jersey.

Eero, M., H. V. Strehlow, C. M. Adams, and M. Vinther. 2014. Does recreational catch impact the TAC for commercial fisheries? ICES Journal of Marine Science 72:450–457.

Fisher, M. R. 1996. Estimating the effect of nonresponse bias on angler surveys. Transactions of the American Fisheries Society 125:118–126.

Fisher, M. R. 1997. Segmentation of the angler population by catch preference, participation, and experience: a management-oriented application of recreation specialization. North American Journal of Fisheries Management 17:1–10.

Gigliotti, L., and A. Dietsch. 2014. Does age matter? The influence of age on response rates in a mixed-mode survey. Human Dimensions of Wildlife 19:280–287.

Harrell, F. E. Jr., and with contributions from C. Dupont and many others. 2016. Hmisc: Harrell miscellaneous. R package version 4.0-2. R Foundation for Statistical Computing, Vienna. Available: http://cran.r-project.org/package=Hmisc. (May 2017).

Hartill, B. W., M. Cryer, J. M. Lyle, E. B. Rees, K. L. Ryan, A. S. Steffe, S. M. Taylor, L. West, and B. S. Wise. 2012. Scale- and context-dependent selection of recreational harvest estimation methods: the Australasian experience. North American Journal of Fisheries Management 32:109–123.

Hartill, B. W., and C. T. T. Edwards. 2015. Comparison of recreational harvest estimates provided by onsite and offsite surveys: detecting bias and corroborating estimates. Canadian Journal of Fisheries and Aquatic Sciences 72:1379–1389.

Holt, D., and T. M. F. Smith. 1979. Post stratification. Journal of the Royal Statistical Society Series A (General) 142:33–46.

Hutt, C. P., and P. W. Bettoli. 2007. Preferences, specialization, and management attitudes of trout anglers fishing in Tennessee tailwaters. North American Journal of Fisheries Management 27:1257–1267.

Kelley, K. 2007. Sample size planning for the coefficient of variation from the accuracy in parameter estimation approach. Behavior Research Methods 39:755–766.

Kelley, K., S. E. Maxwell, and J. R. Rausch. 2003. Obtaining power or obtaining precision: delineating methods of sample-size planning. Evaluation and the Health Professions 26:258–287.

Kelley, K., and J. R. Rausch. 2006. Sample size planning for the standardized mean difference: accuracy in parameter estimation via narrow confidence intervals. Psychological Methods 11:363–385.

Laborde, L. P., F. C. Rohwer, M. D. Kaller, and L. A. Reynolds. 2014. Surveying Louisiana waterfowl hunters: open web and random mail surveys produce similar responses to attitudinal questions. Wildlife Society Bulletin 38:821–826.

Landsman, S. J., H. J. Wachelka, C. D. Suski, and S. J. Cooke. 2011. Evaluation of the physiology, behaviour, and survival of adult Muskellunge (*Esox masquinongy*) captured and released by specialized anglers. Fisheries Research 110:377–386.

Larkin, M. F., J. S. Ault, R. Humston, and L. Lou. 2010. A mail survey to estimate the fishery dynamics of southern Florida's Bonefish charter fleet. Fisheries Management and Ecology 17:254–261.

Lenth, R. V. 2001. Some practical guidelines for effective sample size determination. American Statistician 55:187–193.

Lewis, E. F., M. Hardy, and B. Snaith. 2013. Estimating the effect of nonresponse bias in a survey of hospital organizations. Evaluation and the Health Professions 36:330–351.

Lukacs, P. M. 2007. Development and analysis of joint internet–telephone hunter surveys. Human Dimensions of Wildlife 12:263–273.

Maunder, M. N., and A. E. Punt. 2004. Standardizing catch and effort data: a review of recent approaches. Fisheries Research 70:141–159.

McCormick, J. L., D. Whitney, D. J. Schill, and M. C. Quist. 2015. Evaluation of angler reporting accuracy in an off-site survey to estimate statewide steelhead harvest. Fisheries Management and Ecology 22:134–142.

Peress, M. 2010. Correcting for survey nonresponse using variable response propensity. Journal of the American Statistical Association 105:1418–1430.

Peterman, R. M. 1990. Statistical power analysis can improve fisheries research and management. Canadian Journal of Fisheries and Aquatic Sciences 47:2–15.

R Development Core Team. 2014. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.

Roach, B., J. Trial, and K. Boyle. 1999. Comparing 1994 angler catch and harvest rates from on-site and mail surveys on selected Maine lakes. North American Journal of Fisheries Management 19:203–208.

Ruxton, G. D. 2006. The unequal variance *t*-test is an underused alternative to Student's *t*-test and the Mann–Whitney *U*-test. Behavioral Ecology 17:688–690.

Peytchev, A. 2013. Consequences of survey nonresponse. Annals of the American Academy of Political and Social Science 645:88–111.

Vannieuwenhuyze, J., and G. Loosveldt. 2013. Evaluating relative mode effects in mixed-mode surveys: three methods to disentangle selection and measurement effects. Sociological Methods and Research 42:82–104.

Vannieuwenhuyze, J., G. Loosveldt, and G. Molenberghs. 2010. A method for evaluating mode effects in mixed-mode surveys. Public Opinion Quarterly 74:1027–1045.

Vaske, J. J. 2002. Communicating judgments about practical significance: effect size, confidence intervals and odds ratios. Human Dimensions of Wildlife 7:287–300.

Wallen, K. E., A. C. Landon, G. T. Kyle, and M. A. Schuett. 2016. Mode effect and response rate issues in mixed-mode survey research: implications for recreational fisheries management. Fisheries Research 36:852–863.

Walther, B. A., and J. L. Moore. 2005. The concepts of bias, precision and accuracy, and their use in testing the performance of species richness estimators, with a literature review of estimator performance. Ecography 28:815–829.

Wickham, H. 2009. ggplot2: elegant graphics for data analysis. Available: http://ggplot2.org. (May 2017).

Wickham, H., and R. Francois. 2016. dplyr: a grammar of data manipulation. R package version 0.5.0. R Foundation for Statistical Computing, Vienna. Available: http://cran.r-project.org/package=dplyr. (May 2017).

Zarauz, L., J. Ruiz, A. Urtizberea, E. Andonegi, E. Mugerza, and I. Artetxe. 2015. Comparing different survey methods to estimate European Sea Bass recreational catches in the Basque Country. ICES Journal of Marine Science 72:1181–1191.

Zimmerman, D. W. 2012. A note on consistency of non-parametric rank tests and related rank transformations. British Journal of Mathematical and Statistical Psychology 65:122–144.