Fostering Sustainable Recreational Fisheries Through Informed Management Decisions: A Revamped Approach for Collecting Data from White Sturgeon Anglers

by

Boyd Nkosane Barrett

B.Sc., University of Technology Jamaica, 2008

Project Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Resource Management

Report No. 640

in the
School of Resource and Environmental Management
Faculty of Environment

SIMON FRASER UNIVERSITY

Spring 2016

Approval

Name:	**Boyd Nkosane Barrett**
Degree:	**Master of Resource Management**
Report No:	**640**
Title:	***Fostering Sustainable Recreational Fisheries through Informed Management Decisions: A Revamped Approach for Collecting Data from White Sturgeon Anglers.***
Examining Committee:	**Chair:** Tjaša Demšar Master of Resource Management Candidate

Andrew Cooper
Senior Supervisor
Associate Professor

Brett van Poorten
Supervisor
Adjunct Professor

Date Defended/Approved: July 12, 2016

Ethics Statement

The author, whose name appears on the title page of this work, has obtained, for the research described in this work, either:

a. human research ethics approval from the Simon Fraser University Office of Research Ethics

or

b. advance approval of the animal care protocol from the University Animal Care Committee of Simon Fraser University

or has conducted the research

c. as a co-investigator, collaborator, or research assistant in a research project approved in advance.

A copy of the approval letter has been filed with the Theses Office of the University Library at the time of submission of this thesis or project.

The original application for approval and letter of approval are filed with the relevant offices. Inquiries may be directed to those authorities.

Simon Fraser University Library
Burnaby, British Columbia, Canada

Update Spring 2016

Abstract

Off-site recreational fishery surveys, when compared to on-site surveys, allows fisheries managers to contact a larger sample over a wider spatial scale at a lower cost. However, off-site surveys are prone to nonresponse bias. Nonresponse bias is known to have adverse effects on sample estimates and can erode the leverage of benefits provided by off-site surveys. I explored nonresponse bias in an off-site survey administered to estimate annual total effort and catch in British Columbia's lower and middle Fraser River white sturgeon (*Acipenser transmontanus*) recreational fishery. I explored biases associated with survey mode and response rate. I further used simulation modeling to determine how sample size affects both survey costs and estimates' accuracy. I found that nonresponse bias arose from anglers' participation rate and to a lesser extent from anglers' catch. Anglers who did not fish were less likely to respond. Simulation modelling showed that sample size in the first phase of contact could be reduced by 40%, while holding the follow-up contact at current sample size, and still produce accurate results. Generally, results show that nonresponse bias affected off-site survey estimates even in a relatively small group of specialized anglers.

Keywords: Off-site angler survey; nonresponse bias; white sturgeon; resampling; follow-up survey, sensitivity analysis

Acknowledgements

I would like to express my sincere gratitude to Dr. Wolfgang Haider, Dr. Andrew Cooper and Dr. Brett van Poorten for their contribution to this project and my education. Thank you to Duane Jesson for providing feedback throughout this project. Also, thanks to my family for their support.

The Ministry of Forests, Lands and Natural Resource Operations (FLNRO) provided data for this project. The principal investigator of this project recognizes the Habitat Conservation Trust Foundation and 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.

HABITAT
CONSERVATION TRUST
FOUNDATION

Table of Contents

List of Tables

List of Figures

1 Introduction

1.1 Introduction

Collecting data from recreational anglers is expensive, labor-intensive, and difficult, due to dispersed angler populations and large management areas (American Fisheries Society, 2011; National Research Council, 2006). In response to the aforementioned challenges, fisheries managers often use off-site angler surveys (i.e., via mail, email, and telephone), as opposed to on-site surveys (e.g., creel surveys), to collect data from dispersed angler populations. Off-site angler surveys provide two advantages compared to on-site surveys; namely: the ability to contact a larger sample over a wider spatial scale, and lower cost per contact (Hartill & Edwards, 2015; Zarauz et al., 2015). However, despite these benefits, off-site angler surveys are prone to nonresponse bias (Mccormick, Whitney, Schill, & Quist, 2015). Nonresponse bias is of growing concern due to downward trends in survey response rates and potential adverse effects that a low response rate has on survey estimates (Brick & Williams, 2012; Fowler, 2013). While response rate maximizing strategies (see for example Dillman, Smyth, & Christian, 2014) are available, they are typically expensive and might not be feasible at larger sample sizes. If response rate maximizing strategies are to be used, fisheries managers often have to make a trade-off between sample size and data quality. This research uses data from a recreational fishery to explore the effect of nonresponse bias on population parameters and also show how simulation modeling can be used to aid in making decisions about the trade-off between sample size, and accuracy and precision of estimates.

Nonresponse bias occurs when “a significant number of people in the survey sample do not respond to the questionnaire and are different from those who do in a way that is important to the study” (Connelly, Brown, & Knuth, 2000). Generally, nonresponse bias is a function of response rate and the difference between respondents and nonrespondents with regard to the survey variables of interest (Marsden & Wright, 2010). More specific to recreational fishery surveys, nonresponse bias often occurs when respondents are avid anglers who have greater experience, more successful fishing trips, and expend more effort on the fishery than nonrespondents (Zarauz et al., 2015). Avid anglers are more likely to participate in recreational fisheries surveys (Fisher, 1996); although this is not always the case (Larkin, Ault, Humston, & Lou, 2010). However, if avid anglers are indeed overrepresented amongst respondents, estimates unadjusted for nonresponse bias will have a deleterious effect on population parameters (Fisher, 1996). Nonresponse bias tends to result in overestimated population parameters in off-site angler surveys (Connelly et al., 2000; Fisher, 1996). Overestimates of fishing effort to the magnitude of 25% has been reported (Connelly et al., 2000).

Surveys with a low response rate breach the fundamental concept of collecting a random sample (Brick, 2013), and their results should be viewed with skepticism if a nonresponse bias assessment is not undertaken (Lewis, Hardy, & Snaith, 2013). Interestingly, nonresponse bias has even been observed in surveys with moderately high (i.e., 62% response rate) response rate (Fisher, 1996). It is therefore important to test survey estimates for the effect of nonresponse bias, especially when response rate is low (Fisher, 1996). But first, attempts must be made to increase response rate since a high response rate reduces the potential for response bias (Lew, Himes-Cornell, & Lee, 2015).

There are three major themes in survey research that focus on increasing survey response rate and minimizing nonresponse bias (Brick, 2013). These include studies that: (1) explore the theoretical aspect of the response mechanisms causing a member to respond; (2) explore data collection mechanisms that increase response rates; and (3) identify statistical methods to correct nonresponse bias. I will briefly explore each of these below.

There are several studies that propose theories to explain the sociological or psychological response mechanisms that cause a sample member to respond. One example, leverage-saliency theory (Groves, Singer, & Corning, 2000), states that individuals place varying degrees of importance on survey attributes. Sample members will weigh (i.e., cost/benefits) these attributes to determine if they will participate. Surveys should therefore incorporate a variety of attributes to attract a diverse set of respondents (Groves, Presser, & Dipko, 2004). These attributes include importance of survey topic, incentives, and other benefits of participating. Making a single attribute salient (for example, making survey topic the only salient attribute) will attract a homogenous group of respondents and likely result in nonresponse bias (Groves et al., 2004).

There are numerous data collection methods that can aid in increasing response rates. An example, the Tailored Design Method (see Dillman, 2011), provides guidance on survey design and implementation. Other authors have suggested using multiple survey modes to increase response rate (Dillman et al., 2014; Zarauz et al., 2015), double-sampling or two-phase sampling (Valliant, Dever, & Kreuter, 2013), together with the use of responsive designs (Groves & Heeringa, 2006). The objective of each of these methods is to maximize response rate and diversify the group of respondents through additional survey effort.

Once response rates are maximized, it is still important to employ statistical methods to correct for any nonresponse bias. These methods include post-survey adjustment strategies such as weighting and data imputation methods (see for example: Lew et al., 2015; Vaske, Jacobs, Sijtsma, & Beaman, 2011). Post-survey adjustment strategies are used to correct for nonresponse bias and sample misrepresentation in surveys with missing data (Peytchev, 2013; Singer & Ye, 2012). However, these adjustment methods require auxiliary data with a strong effect on variable outcome. The use of auxiliary variables with a weak effect on variable outcome will increase the variance of estimates without reducing bias (Krueger & West, 2014).

Strategies aimed at increasing survey response rate and reducing the effect of nonresponse bias are often necessary but can be expensive (Peytchev, 2013). Surveys must be optimally designed if these strategies are to be incorporated into data collection without incurring extreme additional costs. Optimally designed surveys aim to minimize cost while achieving acceptable precision, accuracy and power. Authors in other fisheries related studies have used simulation modeling to test the accuracy and precision of estimates based on varied levels of survey effort. For example, Barnes et al. (2014) used simulation modeling to show that effort, catch, and harvest estimates were still reliable even when survey effort was reduced to 33% of the original survey. Also, Xu et al. (2015) used post-survey simulation to improve future survey design by showing that estimates were highly accurate even when sampling effort was reduced.

The objective of this thesis is to explore nonresponse bias in a standard survey administered to evaluate effort and catch estimates in the white sturgeon (*Acipenser transmontanus*) fishery in the lower Fraser River, British Columbia. I explore bias associated with survey mode and response rate. I further use simulation modeling to determine how sample size affects both survey costs and accuracy of estimates of catch and effort.

1.2 Background and Study System

There are six recognized populations of white sturgeon (*Acipenser transmontanus*) in British Columbia. Four populations are legally listed under the Species at Risk Act (SARA), Schedule 1. The other two populations were assessed as endangered by the Committee on the Status of Endangered Wildlife in Canada (COSEWIC). The two endangered populations (i.e., the lower and middle Fraser River populations) are subjected to a catch-and-release fishery (Committee on the Status of Endangered Wildlife in Canada, 2012). This fishery is known as the Lower Fraser River White Sturgeon Recreational Fishery. The fished area encompasses the Harrison River, Pitt River and eight other management regions that extend from the Mission Bridge to Williams Lake (Figure 1).

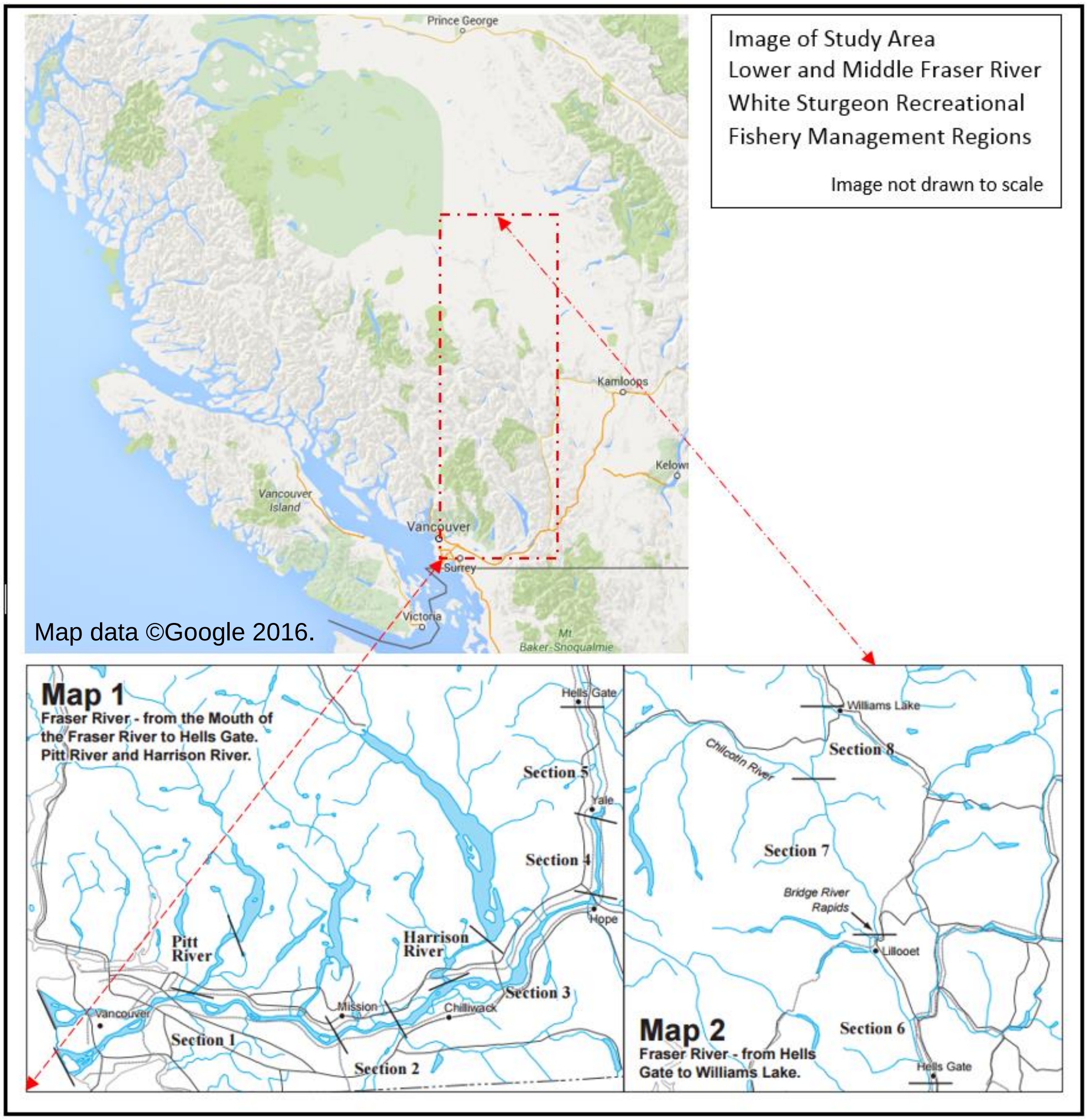


Figure 1. Image of study area

Note: The red box indicates the general location of the study area. The image is not drawn to scale. Sources: Upper image – Map data **© Google 2016.** Map 1 and Map 2 were adapted from English & Jesson (2013).

Anglers who participate in the Lower Fraser River White Sturgeon Recreational Fishery are required to purchase a White Sturgeon Conservation Licence (WSCL) in addition to a basic provincial angling licence (Ministry of Forests Lands and Natural Resource Operations, 2015). The WSCL is valid for one fishing season (between April 1 and March 31 the following year). Anglers can purchase a 1-Day, 8-Day or an Annual licence. Holders of a WSCL either fish with the assistance of a licensed white sturgeon guide (these anglers are called “guided anglers” within this project) or without a licensed guide (these anglers are called “non-guided anglers” within this project).

The Annual Fraser River White Sturgeon Angling Questionnaire (referred to as the “sturgeon angler survey” within this project) is the official survey used to collect data from WSCL holders. The Ministry of Forests, Lands and Natural Resource Operations (FLNRO) administer the sturgeon angler survey. Data obtained from the sturgeon angler survey are used to calculate the catch-per-unit-effort (CPUE) for non-guided anglers. The sturgeon angler survey began in 2009 (English & Jesson, 2013) and was modified for the 2013/14 survey based on recommendations made by English and Jesson (2013). Key features of the original (2009/10 to 2012/13) survey included: (1) voluntary participation; (2) census of all licensees; (3) the survey was disseminated via mail; (4) anglers were asked to report how they fished (guided, non-guided or did not fish); and (5) non-guided anglers were asked to report their effort (number of days fished) and catch (number of white sturgeon caught and released) for each management region (English & Jesson, 2013). Modifications to the survey included (1) selecting a random sample of anglers to participate in the survey and (2) incorporating an email survey, reminder notifications, and telephone follow-up survey of nonrespondents into the survey methodology. The sturgeon angler survey was changed to a two-phase sampling methodology.

The recommendations made by English and Jesson (2013) were primarily geared towards reducing nonresponse bias, reducing survey cost, and improving survey efficiency. This research aims at using the results of the 2014/2015 sturgeon angler survey to do an assessment of the effect that bias associated with survey mode and nonresponse has on survey estimates. In light of the shift from a census to sample survey, this research aims at identifying an optimum sample size that will achieve an acceptable trade-off between survey cost and accuracy.

2 Literature Review

This section presents literature on social science survey errors, post survey adjustment strategies and resampling methods. The section begins by briefly exploring sampling error, coverage error, and measurement error before providing a more detailed review of nonresponse error. Emphasis is placed on examples from recreational fishery surveys where possible. The final two sections focus on post survey adjustment strategies and resampling methods.

2.1 Social science survey errors

Errors are an inevitable aspect of social science surveys and their occurrence can either be random (i.e., variance) or systematic (i.e., bias) (Kent, 2001). Survey research is primarily focused on systematic survey errors since they can confound survey results. However, it is important to limit random error as well since it limits the power of statistical tests to detect differences between treatments. The literature on survey research identifies four general types of survey errors; these are sampling error, coverage error, measurement error, and nonresponse error (Dillman et al., 2014).

2.1.1 Sampling error

Sampling error refers to “the amount by which the selected sample findings deviates from data that would have been obtained if the entire population were surveyed” (McNabb, 2013, p. 30). Sampling error is a random error and is a direct result of selecting a random sample as opposed to doing a census. Sampling error and sample size are inversely proportional (Dillman et al., 2014).

2.1.2 Coverage error

Coverage error occurs when members of the population of interest are excluded from the sample and is different from those who are included with regard to important survey variables (Dillman et al., 2014). This exclusion can be the result of errors in population specification, false or inaccurate selection procedures (e.g., convenience vs. probability sample), and issues resulting from using an inaccurate or incomplete sample frame (McNabb, 2013). Dillman et al. (2014) also pointed out that coverage error can result from the use of a survey mode that all sample members do not have access to.

2.1.3 Measurement error

Measurement error arises when participants provide incorrect or imprecise answers (Dillman et al., 2014). This includes “the return of false or subjectively modified information from survey respondents” (McNabb, 2013, p. 114). Angler surveys are prone to measurement errors since they rely on angler self-reporting data; this is especially true for off-site angler surveys. Anglers are more likely to overestimate values as the frequency of participation and recall period increases (Connelly & Brown, 2011). Measurement error and the resulting bias are manifested in four different forms; these are: (1) recall bias, (2) reporting bias, (3) prestige bias, and (4) digit preference. I will briefly explore these biases in the following paragraphs.

2.1.3.1 Recall bias

Recall bias occurs when anglers misreport their data due to a failure to remember required details of the event of interest (Zarauz et al., 2015). Telescoping (i.e., reporting of events that occurred in a different time period than the period in question) and recall decay (i.e., outright failure to recall the required event) can cause recall bias (Connelly et al., 2000).

The effect of recall bias on recreational fishery survey estimates is documented in the survey literature, however the findings are contrary. A common finding among researchers is that participants are more likely to overestimate estimates as recall period increased (Connelly et al., 2000; Tarrant, Manfredo, Bayley, & Hess, 1993). However, Osborn and Matlock (2010) reported that the effect of recall varied based on survey questions and that there was no significant difference in the mean number of days fished based on recall period. Connelly and Brown (2011) compared data from an annual mail survey to data from a three-phase survey (3 mailings over a one year period). The researchers reported that effort estimates from the three-phase survey were significantly different from estimates from the one-year recall survey. However, there was no identifiable consistent pattern in the direction of difference across multiple lakes. Some lakes had a larger value for the three-phase survey while other lakes had a larger value for the annual survey. Recall bias is typically low for off-site surveys if species are rare and catch is memorable (Pollock, Jones, & Brown, 1994). This line of reasoning is sometimes used to justify accuracy of estimates in particularly impressive fisheries (e.g., Mccormick et al., 2015; Pollock et al., 1994).

2.1.3.2 Reporting Bias

Reporting bias is associated with intentionally providing incorrect data. An example of reporting bias was found by McCormick, Quist and Schill (2013), who explored the effect of reporting bias on catch and harvest estimates of Chinook salmon (*Oncorhynchus tshawytscha*) in Idaho. Although no conclusive evidence was provided, McCormick et al. (2013) suggested that anglers underreported catch in an attempt to extend the fishing season. Management of the Chinook fishery is characterized by: (1) closure of the fishery once the harvest share is reached, and (2) a limit on the amount of wild Chinook that are caught accidentally. Managers rely on absolute values of catch and harvest to manage the fishery and the fishery is closed once the quota is reached. Anglers are therefore likely to underreport catch in an attempt to extend the season (McCormick et al., 2013).

In another study, McCormick et al. (2015) assessed the magnitude of reporting bias in catch data reported by anglers in Idaho's state wide off-site steelhead angler survey. The authors compared data reported on anglers' harvest permits to data reported by the same group of anglers in an off-site survey. The authors assumed that estimates from the harvest permits were the true estimates. The authors did not find any significantly consistent pattern of misreporting; however, anglers tended to overestimate catch at lower catch rate and underestimate at higher catch rates.

2.1.3.3 Prestige bias

Prestige bias occurs when anglers provide responses that they deem to be in line with acceptable social norms or boosts self-image (Hartill & Edwards, 2015). Exaggeration of catch rate is an example of prestige bias. Anglers tend to exaggerate catch at lower catch rate to avoid admitting that they did not catch any fish (Sullivan, 2003). Hartill and Edwards (2015) reported prestige bias resulted in overestimates of total catch because unsuccessful anglers reported catching fish and successful anglers over reported catch. Exaggerating catch at low catch rates can mask fish population decline (Sullivan, 2003); this can be detrimental for fish stocks.

2.1.3.4 Digit preference

Participants of recreational surveys utilize a variety of cognitive processes when providing numeric responses to frequency and quantity questions based on memory recall (Vaske & Beaman, 2006). Cognitive processes include multiplicative rule-based decision-making, episode enumeration, digit preference (or number preference), and the use of prototypes (Vaske & Beaman, 2006).

Multiplicative rule-based decision-making is used when events are numerous and similar (Vaske, Huan, & Beaman, 2003). Under multiplicative rule-based recollection, participants develop a frequency rule (e.g., “I fished 3 times per week”), which is then multiplied by the time period in question. Bias will arise when participants fail to adjust estimates for periods when no fishing occurred or when catch rate was lower due to seasonal variation in catch (Vaske et al., 2003). Vaske et al. (2003) stated that the use of multiplicative rule-based decision-making resulted in estimates that were 45% higher than quantity estimates not based on multiples.

In contrast, participants often use episode enumeration when the number of events (frequency questions) or the number of fish caught (quantity questions) to be reported is low (Sudman, Bradburn, & Schwarz, 1996). Episode enumeration entails summing the value for each event based on long-term recollection and providing an answer. Since the events are few and do not occur frequently, responses are more likely to be affected by episode omission (failure to recall specific events) and telescoping (reporting events which occurred in a different time period); less memorable events are more likely to be forgotten (Vaske et al., 2003).

Digit preference occurs when participants provide numeric response ending in rounded values, for example, 0 or 5 (Tarrant et al., 1993). The preference for certain digits changes based on the specific survey questions. For example, participants are likely to provide 4 or 8 when duration is in hours of participation during a single day, and 7, 14, 30, and 60 when participation is measured in days (Vaske & Beaman, 2006).

The process of using prototypes entails the use of a single value to represent a range of possible values. The magnitude of bias resulting from the use of prototypes can be significant since participants might use a prototype to represent a different range of potential values. For example, one participant might use 10 to represent a range of 8 to 12, while another might use 10 to represent a range of 6 to 14. Beaman, Vaske, Schmidt, and Huan (2015) showed that the range of values that a prototype represents increases as the magnitude of responses values increases.

2.1.4 Nonresponse error

Nonresponse error is a non-sampling error (Assael & Keon, 1982) and occurs when respondents differ from nonrespondents with regard to important survey variables (Connelly et al., 2000). The literature reviewed in this section focuses on unit nonresponse (i.e., no response from members of the sample) as opposed to item nonresponse (i.e., no response on a single question) (Littvay, Popa, & Fazekas, 2013). This research defines a nonrespondent as a member of a sample who is selected to participate but for which data are not obtained for any reason, excluding *no successful contact* (Carkin & Tracy, 2015). This research also accepts the fact that data from nonrespondents will never be known (Fowler, 2013); it can only be estimated (Peress, 2010).

The response rate of a survey is not a true indicator of the representativeness of a sample, or the presence of nonresponse error (Assael & Keon, 1982). Therefore, response rate should not be used as the sole measure of survey success (Fuchs, Bossert, & Stukowski, 2013). Peytchev (2013) adopted a formula from Groves (1989) and showed that nonresponse bias, being a surrogate measure of nonresponse error, is a function of both response rate and the difference between the estimated mean value for respondents and nonrespondents. This difference pertains to important survey variables. When a low response rate is obtained, simply extrapolating estimates from respondents to the sample and the population is questionable (Fisher, 1996). Unit nonresponse can make the sample non-representative of the population, and when ignored in the data analysis phase, will likely lead to biased results (Lew et al., 2015). Peytchev (2013) highlighted the potential effect of nonresponse bias on estimates of means, proportions, variance, and correlations. Regarding the mean, Peytchev (2013) stated that nonresponse bias could lead to serious over- or underestimates. Nonresponse bias commonly leads to an underestimation of the variance, however, there is not much evidence of the effect of nonresponse bias on the correlation between variables (Peytchev, 2013). Brick (2013) goes as far as stating that that a low response rate breaches the fundamental assumption of most statistical analysis procedures, namely, a random sample; since respondents a self-selected into the sample.

As noted above, nonresponse bias is a function of response rate and the difference between respondents and nonrespondents. Within recreational fishing, anglers are believed to fall along a continuum that range from novice to avid—highly experienced anglers. This concept is called angler specialization (Fedler & Ditton, 1994). An angler's catch rate, catch motivation, and perception of resource conservation will vary based on where they fall along the continuum. Avid anglers are characterized as being frequent and committed participants to a particular fishery (Needham, Scott, & Vaske, 2013), more vocal about the species they fish for—compared to the novice anglers—(Beardmore, Hunt, Haider, Dorow, & Arlinghaus, 2014), and have honed their skillset for catching a particular species (Johnston, Arlinghaus, & Dieckmann, 2010). Avidity bias occurs when there is a disproportionate representation of avid anglers amongst respondents. Avidity bias will arise if avid anglers are more likely to respond to a survey compared to novice angles (Thomson, 1991). Thompson (1991) noted that avidity bias is commonly associated with sampling anglers in creel surveys, but the concept can also be extended to off-site angler surveys. This misrepresentation of avid anglers amongst respondents is a key source of nonresponse bias (Thomson, 1991).

The effect of nonresponse bias on recreational survey fishing estimates is documented in the literature. Fisher (1996) explored the effect of nonresponse bias on catch estimates derived from an off-site angler survey by assessing the difference between estimates that were corrected for nonresponse error (using nonresponse post survey correction methods) and estimates that were not adjusted. Fisher (1996) noted that: nonresponse bias caused certain population subgroups to be misrepresented—resulting in biased estimates—and that early (i.e., fishers who responded fastest to the survey) respondents had higher participation rates and years of experience than anglers in the general population. In a more recent study, Zarauz et al. (2015)reported that evidence suggest that nonrespondents to a mail survey of recreational anglers were less experienced and had less successful fishing trips than respondents. Lew et al. (2015), in a survey of charter boat owners, found that nonrespondents were more likely to fish late in the season and during the off-season when catch was lower.

As opposed to the authors in the preceding paragraph, Larkin et al. (2010) did not find any statistically significant difference between nonrespondents and respondents. More specifically, the authors found that respondents and nonrespondents did not differ in years in the fishery, effort, or mean size of bonefish (*Albula vulpes*) caught in the previous year. While the survey had a 59% response rate, Larkin et al. (2010) noted that his research might have lacked statistical power. Notwithstanding this limitation, Larkin's work provided evidence that nonresponse bias is not always present in face of a low response rate if anglers are homogeneous with respect to the survey variable of interest.

In general, research focused on nonresponse aims to propose methods to increase survey response rate and use statistical procedures to correct biased estimates (Brick, 2013). The following sections will explore the existing literature on methods available to increase response rate and statistical procedures used to correct biased estimates.

2.2 Increasing survey response rate

The literature provides a diverse array of methods to increase the response rate of social science surveys (see for example Dillman et al., 2014). Two methods that are prominently featured and that are intricately linked to this research are the use of mixed-mode surveys and follow-up contact (two-phase sampling). Mixed mode surveys refer to the combination of different survey modes of contact sample members. For example, mail survey combined with an email survey, or with a telephone survey (Dillman et al., 2014). Mixed mode surveys can diversify the pool of respondents (Vaske et al., 2011) and reduce the occurrence of differential response rate (Gigliotti & Dietsch, 2014). Differential response rate occurs when population subgroups respond disproportionately to a survey. For example, age-related differential response rate occurs when people of different age groups respond differently to a survey based on survey characteristics (Gigliotti & Dietsch, 2014). Differential response rate can be the effect of age, sex, economic status, and other characteristics of the population. Nonresponse bias will arise if survey modes collect data disproportionally and members of these population subgroups differ in respect to survey variables of interest (Gigliotti & Dietsch, 2014).

Differential response rate has a direct impact on recreational fishing survey estimates. Recently, Zarauz et al. (2015) investigated the effect of differential response rate on estimates obtained from the Basque Country's sea bass recreational fishery. Data were collected via regular mail, email, and telephone. Participants were asked to report their age, experience (number of years fishing), total effort (days), and total catch (in kg). Zarauz et al. (2015) reported that each survey mode yielded different estimates. For example, total catch estimates for shore anglers were 129, 156, and 351 tonnes for email, phone, and mail surveys, respectively. The authors attributed this difference to differential response rate. The authors also stated that recall bias likely had an effect on estimates. In another study, Laborde, Rohwer, Kaller, and Reynolds (2014) compared the results of a random mail survey and an internet convenience sample. The surveys sought to obtain data on waterfowl-hunting effort, success, satisfaction, regulatory alternatives, and demographics. Laborde et al. (2014) reported that respondents to the internet survey hunted more often, harvested more waterfowls, and placed greater importance on waterfowl hunting. Therefore, the use of a single survey mode can yield incorrect estimates (Laborde et al., 2014; Zarauz et al., 2015).

Two-phase sampling or double-sampling entails the identification and contact of a subsample from the original sample. There is a fundamental difference between double-sampling for stratification and double sampling for nonresponse (Valliant et al., 2013). Double sampling for nonresponse, as proposed by Hansen & Hurwitz (1946), entails the random selection of a sample from a population; subsequently, the sample is divided into two strata (i.e., respondents and nonrespondents). A second random sample is then selected from the group of nonrespondents. Double-sampling for nonresponse is built on the premise that the sample taken from the nonrespondents is a true representation of all nonrespondents (Lohr, 2010). Therefore, a high response rate is necessary in the follow-up contact. Double-sampling usually requires more intensive survey methods than was used in the first contact and has been criticized for being expensive (Thompson, 2012).

Using mixed mode surveys and double sampling can potentially increase response rate and reduce response bias; however, a 100% response rate is rarely attained (Peytchev, 2013). In response, authors have advocated for the use of post-survey adjustment strategies to adjust estimates obtained from surveys. Post survey adjustment strategies will be explored in the next section.

2.3 Adjusting estimates for nonresponse

In the case where a simple random sample is used to select sample members, and all members of the sample respond, the sample estimates can be used as an unbiased estimator of the population mean (Lew et al., 2015). However, in the case where data are missing due to unit nonresponse, simply extrapolating data from respondents to the population can result in incorrect estimates (Fisher, 1996; Lew et al., 2015). Weighting class adjustment, poststratification, and raking adjustment are commonly used to adjust population estimates in cases where 100% response rate is not obtained and auxiliary data are available for the population. Weighting class adjustment and poststratification are briefly explored here.

Survey data that are missing can be classified as being missing completely at random, missing at random, or as nonignorable nonresponse (Lohr, 2010). Missing completely at random refers to a situation where the likelihood of responding is not the result of the survey variable being investigated or auxiliary variables. No correction is needed in this case. Missing at random refers to a situation where the probability of response does not depend on the survey variable being investigated but is related to the auxiliary variables (Lohr, 2010). Auxiliary data can be used to create a model for response in the case of data that are missing at random (Valliant et al., 2013). Finally, nonignorable response occurs when the probability of responding depends on the survey variables being investigated and dependence cannot be eliminated by modeling responses based on auxiliary variables. Valliant et al. (2013) pointed out that the problem with this approach is that data are not available for nonrespondents; therefore modeling is impracticable. Sample weighting can only be used in the case of data that are missing at random (Valliant et al., 2013).

Weighting class adjustment is used to correct sample estimates for non-sampling errors (Lohr, 2010). Weighting class adjustment methods use data available about all sample members to form weighting classes (Lew et al., 2015). Weighting class variables often include demographic data or other variables that are important to the survey (Lohr, 2010). Similar sample members are grouped in a class and an explicit assumption is made that respondents in a class are a perfect representation of nonrespondents in the same class. Respondents are then assigned an increased weight to compensate for nonrespondents (Lohr, 2010).

Fisher (1996) used weighting class adjustment to correct estimates obtained from an angler survey with a 62% response rate. Firstly, Fisher (1996) used logistic regression to estimate the response probability of sample members; the independent variables were age, gender, race, and licence purchase date. Additionally, a linear regression was fitted for total effort and years of experience to understand how these variables affected response. Regression residuals were then correlated with response probabilities to determine if the likelihood of responding depended on total effort and years of experience. Anglers were then placed in an adjustment cell based on their response propensity. The mean response propensity of the cell was assigned to each member of the cell. The adjustment weight was then multiplied by the design weight and the product used to calculate all survey results.

Similar to the weighting procedure outlined above, post stratification is used to correct for nonresponse bias and coverage error, reduce variance, and improve the precision of estimates (Lohr, 2010). Poststratification is characterised by the use of population counts to adjust weights. A fundamental difference between poststratification and weighting class adjustment is that the number of units in each stratum in the population is known for poststratification while it is unknown, hence estimated, for weighting class adjustment (Lohr, 2010).

Lew et al. (2015) provided a detailed review of an application of survey weighting and the effect on survey estimates. Lew et al. (2015) used the formula below to assign a weight to each respondent:

$$Individual\ weight\ for\ i(W_i) = \ w_{i1} * w_{i2} * w_{i3}, \tag{1}$$

where the weight given to the i_{th} respondent is denoted by W_i. Additionally, w_{i1}, w_{i2}, and w_{i3} are weights that make different types of adjustments. The base weight is denoted by w_{i1}. This base weight is the inverse of the probability of being selected and is calculated by dividing the population size (N) by the size of the sample (n) ($\frac{N}{n}$), this is also called the inclusion probability. The nonresponse adjustment weight is denoted by w_{i2} and makes adjustments for potential differences between respondents and nonrespondents. The post-stratification weight is denoted by w_{i3}. The post-stratification weight adjusts for incomplete coverage and ensures that the sample conforms to known population totals. Auxiliary data can be used to compare the sample to the population and for developing post-stratification weights (Lew et al., 2015). Valliant et al. (2013) pointed out that post-stratification weights are also useful to reduce standard errors.

2.4 Resampling procedures for recreational fishery surveys

Resampling allows researchers to make inferences by taking repeated samples from the original sample. Four types of resampling are: (1) randomization test, (2) cross-validation, (3) jackknife, and (4) bootstrap (Bai & Pan, 2008). This research used a variation of the bootstrap procedure which can be used to estimate a population parameter θ. B bootstrap samples are generated and an estimate of the population parameter $(\hat{\theta})$ is calculated for each bootstrap iteration, these estimates are denoted $\hat{\theta}^*_1, \ldots, \hat{\theta}^*_B$ (Puth, Neuhäuser, & Ruxton, 2015). Each iteration of the bootstrap has the same sample size as of the original sample, samples are drawn with replacement (Puth et al., 2015), and no assumptions are made about the underlying distribution of the data (Mooney, Duval, & Duval, 1993) The underlying premise of the bootstrap is commonly adapted in resampling simulation test under the common name resampling.

Resampling procedures offer researchers the ability to simulate data collection procedures and obtain the bias, accuracy, and precision of estimates without any additional primary data collection (Barnes, Simpson, Carreiro, & Voorhees, 2014). This research drew upon examples in the literature (e.g., Barnes, 2014; Lowry, Stick, Lindquist, & Cheng, 2015; Xu, Zhang, Xue, Ren, & Chen, 2015). In this context, resampling was used to determine an optimal sample size for fishery related data collection. For example, Lowry et al. (2015) used bootstrap resampling to determine the uncertainty of surf smelt (*Hypomesus pretiosus*) recreational fishery catch estimates based on data obtained from creel surveys undertaken in Washington. Similarly, Barnes et al. (2014) took resamples from 3 years of creel survey data to determine the accuracy of estimates based on three levels of creel survey effort. Barnes et al. (2014) showed that estimates were still representative even at one-third of the original effort; however, confidence intervals were large. Barnes et al. (2014) noted that estimates were accurate enough, given the nature and characteristics of the fishery. The simulation model therefore pointed managers towards an optimal sample size where cost and accuracy needs were met (Barnes et al., 2014). Finally, Xu et al. (2015) used resampling to identify an optimal sample size for a bottom trawl survey and in so doing sought to reduce survey cost and negative environmental effects associated with the survey methodology. The mean squared error (MSE) of estimates at different sample sizes was used to assess optimality. Through resampling, these authors showed that sample size could be drastically reduced while still obtaining precise and accurate data.

2.5 Research questions

The literature provides evidence of the potential negative effects of nonresponse bias on off-site recreational fishing survey estimates when response rate is low (Mccormick et al., 2015). However, these potential negative effects are not always realized since nonresponse bias is a function of response rate and the difference between respondents and nonrespondents (Peytchev, 2013). Therefore, nonresponse bias will be negligible if the group is homogenous with regard to the survey variable of interest (Larkin et al., 2010). Three common themes resonate in the literature on nonresponse bias: (1) a low response rate is not a direct indication of response bias (Fuchs et al., 2013); (2) mixed-mode surveys can be used to diversify the pool of respondents and reduce nonresponse bias (Gigliotti & Dietsch, 2014); and (3) adjustment methods should be used to correct for nonresponse bias when response rate is low (Lew et al., 2015). Based on these three themes, this research asked: *How are catch and effort estimates obtained from The Annual Fraser River White Sturgeon Angling Questionnaire affected by nonresponse bias*?

Secondly, the literature identifies resampling as a valuable tool that enables researchers to optimise future data collection efforts by identifying the optimal sample size needed to obtain a desired level of accuracy and/or precision (see for example Barnes et al., 2014; Lowry et al., 2015; Xu et al., 2015). Based on the conclusions provided by the aforementioned authors, this research will also sought to ask: *How do the accuracy and precision of survey estimates vary as a function of sample size and survey effort?*

3 Methods

This section presents the data collection procedure and the methods of analyses used for this project. The Ministry of Forests, Lands and Natural Resource Operations (FLNRO) collected data used in this research.

3.1 Data Collection

The Annual Fraser River White Sturgeon Angling Questionnaire survey was used to collect data for this project. You can refer to Appendix A for a schematic diagram of the sample selection process that was used here. The population was the entire group of anglers who purchased a WSCL for the 2014/15 white sturgeon fishing season. The population was divided into two mutually exclusive groups, herein called group_1 (Gp1) and group_2 (Gp2). Gp1 consisted of anglers who, at the time of purchasing their WSCS, provided an email address and consented to being contacted via email, Gp2 consisted of all other anglers. Two simple random samples were concurrently selected from Gp1; herein called Gp1_electronic and Gp1_paper. Members of Gp1_electronic were sent a questionnaire via email. Members of Gp1_paper were sent a questionnaire via regular mail. A simple random sample was selected from Gp2 (i.e., Gp2_paper; *please note the difference between Gp1_paper and Gp2_paper*). Members of Gp2_paper were also sent a questionnaire via regular mail. Each sampled angler was sent a single questionnaire regardless of the number of licence he/she purchased.

The survey was a multi-mode survey with three modes: regular mail, email and telephone. The data collection procedure within each group followed a two-phase sample design where a simple random sample was selected for the first contact and a subsample was selected from the group of nonrespondents for the second contact. The regular mail and email modes were used for the first contact and the telephone mode was used for the second contact. Data collection followed Dillman's Tailored Design Survey method (Dillman et al., 2014).

3.1.1 Regular mail survey

The regular mail survey (n = 4,470) consisted of anglers from Gp1_paper ($n_{Gp1.paper}$ = 1,098), and Gp2_paper ($n_{Gp2.paper}$ = 3,372). Data collection procedure for the regular mail survey entailed a first mail-out and a reminder notification. The first invitation to participate was sent on June 15 2015 and a reminder notification was sent 6 weeks later. Each Canadian participant was sent a DL letter sized envelope containing a paper questionnaire and a stamped return envelope. Anglers with an address outside of Canada were required to pay their own return postage. The envelope was delivered by standard mail. Respondents were removed from the list after each contact and only nonrespondents were eligible to be contacted for the nonresponse assessment survey.

3.1.2 Email survey

The email survey ($n_{Gp1.electronic}$ = 4,248) was sent to anglers from Gp1_electronic. Data collection procedure for the email survey entailed a first email contact and a reminder email notification. The first invitation to participate was sent on June 23 2015 and a reminder email was sent 14 days later. Each participant was sent an email requesting participation in the survey. The email had a hyperlink that allowed anglers to connect directly to the survey; responses were automatically stored to a database. Respondents were removed from the list after each contact and only nonrespondents were eligible to be contacted for the nonresponse assessment survey.

3.1.3 Telephone survey

The nonresponse assessment survey was administered via telephone. Due to budgetary constraints, only anglers who provided a Canadian address when purchasing their WSCL were selected to participate. A simple random sample (n = 1,000) was selected from the list of anglers who did not respond to Gp1.electronic ($n_{Gp1.elect.foll}$ = 582), Gp1.paper ($n_{Gp1.paper.foll}$ = 104), and Gp2.paper ($n_{Gp2.paper.foll}$ = 314). Four attempts were made to contact each angler before he or she was deemed as a *no contact*. Calls were made between 5 p.m. and 9 p.m. on weekdays and between 10 a.m. and 8 p.m. on weekends. A professional market research company administered the survey.

3.2 Data Analyses

Only completed questionnaires were used in the analysis. The adjusted response rate was calculated as:

$$Response\ rate\ (\%) = \frac{Number\ of\ completed\ returns}{Total\ sample - Undeliverable\ questionnaires} \times 100 \qquad (2)$$

3.2.1 Statistical Analysis

Chi-square test was used to test for differences in the proportion of licence holders (i.e., 1-day, 8-day, and annual licence holders) in the samples and their respective population. Chi-square test was also used to test for difference in the proportion of anglers (i.e., type of angler – guided, non-guided, and did not fish) in the first contact and follow-up contact. A non-parametric Levene's test was used to test for equality of variance amongst samples. Welch's test, on ranked transformed data, was used to test the null hypothesis of stochastic homogeneity of the variable of interest amongst groups (Ruxton, 2006; Zimmerman, 2012). Dunn's multiple comparison post-hoc test was used for pairwise comparisons when the Welch's test provided evidence of a difference. Type-I error rate (α) was set at 0.05 for all tests.

Annual total catch and annual total effort for the population were only calculated for using

$$\widehat{A_{total}} = \sum_{j=1}^{3} \hat{T}_j, \qquad (3)$$

where $\widehat{A_{total}}$ is the annual total estimate for the entire population and $\hat{T}_j$ is the annual population total estimate for survey j (see Appendix B for a breakdown of calculations.

3.2.2 Sensitivity Analysis

A sensitivity analysis was undertaken to ascertain the uncertainty of estimates as a function of sample size and the number of survey modes. This resampling exercise was executed on the assumption that catch and effort data from the 2014-2015 sturgeon angler survey are a true representation of the population. Three separate sets of resampling were simulated. First, I resampled the original data while varying the sample size only. Second, I resampled the original data while excluding data from the follow-up surveys. Third, I resampled the original data while holding the sample size of the follow-up survey constant. The resampling procedure mimicked that of Hoyle and Cameron (2003) and Zhang et al. (2015) and is outlined below.

1. A total of 15,000 (3 sets × 5 sample size × 1000 iteration per sample size) simple random samples without replacement were selected from the original pool of respondents.
2. The sample size was varied for each iteration (i.e., 100%, 80%, 60%, 40%, and 20%).
3. The annual total catch, annual total effort, and CPUE were calculated for each resample using equation 3.

The costs of the surveys were broken down to cost per angler (total cost / sample size) for the regular mail survey and telephone survey. The email survey had a fixed cost irrespective of sample size.

Root mean square error (RMSE) (Walther & Moore, 2005) was used to evaluate the accuracy of estimates:

$$RMSE = \sqrt{\frac{1}{R}\sum_{i=1}^{R}\left(Y_i^{estimated} - A\right)^2}\ , \qquad (4)$$

where *A* is the true value, based on the original survey; $Y_i^{estimated}$ is the estimated value from the simulation; and *R* is the number of resamples.

4 Results

4.1 General findings

4.1.1.1 Response rates

Gp1_electronic first contact (*n* = 4,248) had the lowest adjusted response rate (22%) amongst all surveys (Figure 2). Interestingly, 63% of respondents to Gp1_electronic follow-up survey stated they did not receive a questionnaire during Gp1_electronic first contact survey. Comparatively, 40% of anglers in Gp1_sample follow-up survey and 44% of anglers in Gp2_paper follow-up survey said that they did not receive a questionnaire during the first contact.

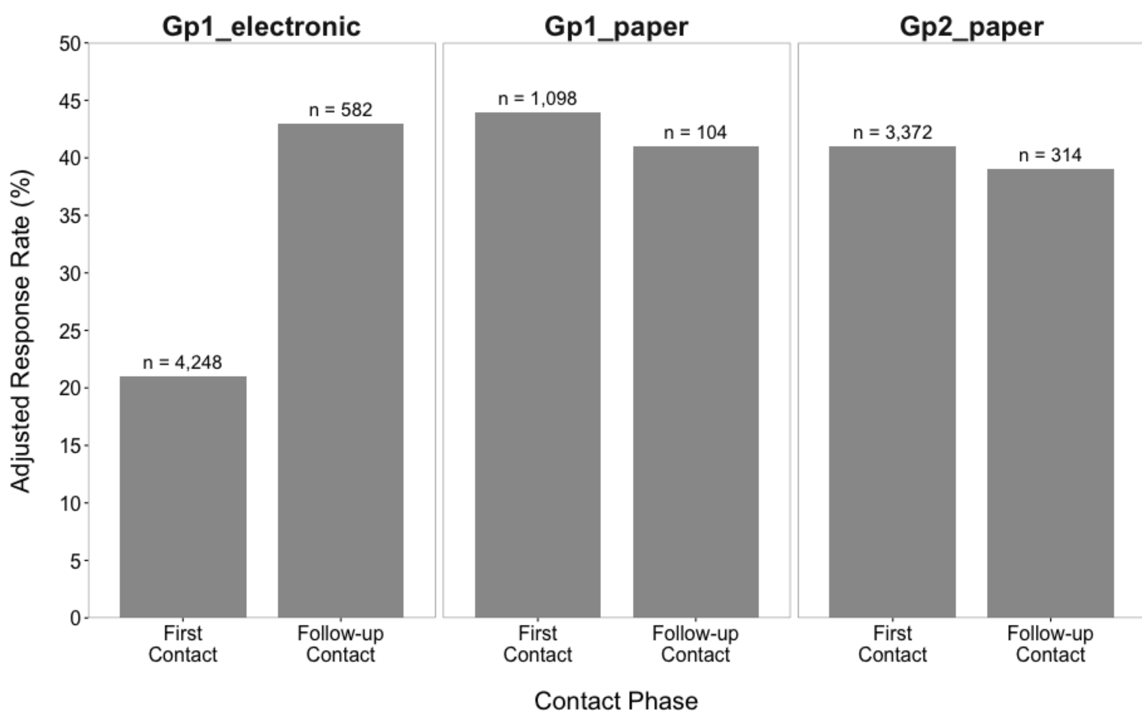


Figure 2. Bar graph showing adjusted response rate for surveys

Note. The adjusted response rate was calculated as [completed questionnaires ÷ (total sent − undeliverable questionnaires)]. Sample sizes are shown above respective bars.

4.1.1.2 Summary statistics

Catch, effort, and CPUE data were right skewed and leptokurtic (see Appendix C). Non-parametric Levene's test computed on the ranked data for survey mode (i.e., email vs. regular mail vs. telephone), group (i.e., Gp1_electronic vs. Gp1_paper vs. Gp2_paper), and phase (i.e., first contact vs. follow-up contact within groups) indicated heterogeneity of variance in all cases.

Twenty-five percent of non-guided anglers accounted for 85% of total effort and 72% of total catch reported for all survey modes. Gp1_paper follow-up had the highest median effort (i.e., 8 days fished), catch (i.e., 6 white sturgeon caught), and CPUE (i.e., 2 white sturgeon caught per angler-day) values (Table 1). Population annual total effort was estimated as 26,023 angler-days, annual total catch was estimated as 61,614 white sturgeon, and CPUE was estimated as 2.4 white sturgeon per angler-day (see Appendix D).

Table 1. Summary statistics for non-guided anglers

Survey Mode	N	Mean	Median	25th percentile	75th percentile	Range
Gp1_electronic						
Effort	252	8.6	4	2	10	119
Catch	252	15.7	4	1	16.5	485
CPUE	252	1.7	1	0.3	2	19.5
Gp1_electronic Follow-up						
Effort	45	6.9	2	1	7	48
Catch	45	13.8	4	0	12	189
CPUE	45	1.9	1	0	3	10
Gp1_paper						
Effort	79	4.8	3	1	3	24
Catch	79	5.3	3	1	6	41
CPUE	79	1.1	1.8	0.3	8	8
Gp1_paper Follow-up						
Effort	13	19.4	8	1	15	99
Catch	13	51	6	3	25	299
CPUE	13	2.2	2	1.7	3.12	4.09
Gp2_paper						
Effort	295	6.1	3	1	6	69
Catch	295	6.7	2	0	6	130
CPUE	295	1.2	0.7	0	1.9	12
Gp2_paper Follow-up						
Effort	29	6.6	2	1	6	59
Catch	29	23.9	1	0	10	500
CPUE	29	1.5	0.6	0	1.7	8.3

Note. The median and the interquartile range (i.e., values at the 25th and 75th percentiles) were used to summarize data because data were non-normally distributed. Range = maximum – minimum value. N is the number of non-guided anglers in the responding sample.

4.2 Survey results

4.2.1.1 Comparing samples to their respective population

Chi-square goodness of fit test for Gp1_electronic provided evidence of a difference between the proportion of licences (i.e. 1-Day, 8-Day and Annual) in the first contact and Gp1's population (X^2 (4) = 26.4, p < 0.0001) (Table 2). Visual inspection of the mosaic plot (see Appendix E) indicated that Annual licensees were overrepresented in the first contact for Gp1_electronic. The chi-square goodness of fit test did not provide evidence of a difference between Gp1_paper, Gp2_paper and their respective populations.

Table 2. Results of chi-square goodness of fit test comparing proportion of licence in samples to their respective population

	Contact Phase											
Group	First Contact vs. Population vs. Follow-up Contact			First Contact vs. Population			Follow-up Contact vs. Population			First Contact vs Follow-up		
	X^2	*df*	*p*	X^2	*df*	*p*	X^2	*df*	*p*	X^2	*df*	*p*
Gp1_electronic	26.4	4	<0.0001*	25.6	2	<0.001*	4.2	2	0.12	7.9	2	0.02*
Gp1_paper	2.1	4	0.7145									
Gp2_paper	3.5	4	0.4777									

Note. An asterisk (*) indicates a significant value (alpha = 0.05).

Non-Canadian anglers were underrepresented in Gp1_electronic (X^2 (2) = 22.4, p < 0.0001). Chi-square test did not provide evidence of a difference for Gp1_paper (X^2 (2) = 1.6, p = 0.45) or Gp2_paper (X^2 (2) = 2, p = 0.37). Non-Canadian anglers were not included in the follow-up survey.

4.2.1.2 Comparing estimates by group: Gp1_electronic, Gp1_paper, and Gp2_paper

At least one group consisted of CPUE estimates that were from a different distribution (Welch's F (2, 224) = 4.8, p = 0.009). A similar pattern was observed for total effort (Welch's F (2, 222.2) = 5.35, p = 0.005) and total catch (Welch's F (2, 222.6) = 8.93, p < .0001). The difference was identified between Gp1_electronic and Gp2_paper for: CPUE ($p = 0.007$), total effort ($p = 0.005$), and total catch ($p < 0.0001$). Dunn's multiple comparison post-hoc test did not provide evidence of a clear difference between Gp1_paper and Gp1_electronic or Gp1_paper and Gp2_paper, in each case. However, the difference between Gp1_paper and Gp2_paper was smaller than the difference between Gp1_paper and Gp1_electronic. Only data from the first contact were used for this analysis.

4.2.1.3 Comparing effort, catch and CPUE estimates for first contact and follow-up within each group

Catch and CPUE estimates from Gp1_paper first contact and Gp1_paper follow-up contact were from different distributions (Welch's F (1, 17.3) = 8.97, $p = 0.008$) and (Welch's F (1, 16.3) = 8.37, $p < .01$), respectively (Table 3). The Welch's test on ranked transformed annual catch data indicated that data from the Gp1_paper follow-up survey ($n = 13$, mean rank = 64.5) tended to be higher than estimates from the first contact ($n = 79$, mean rank = 43.5).

Table 3. Results of the Welch's test comparing ranked estimates from the first contact and follow-up contact for each survey mode

Group	Statistic			
	*df*1	*df*2	F	*p*
Gp1_electronic				
Effort	1	58.47	2.890	0.095
Catch	1	59.10	0.270	0.600
CPUE	1	56.78	0.150	0.700
Gp1_paper				
Effort	1	14.876	2.676	0.123
Catch	1	17.339	8.968	0.008*
CPUE	1	16.312	8.365	0.010*
Gp2_paper				
Effort	1	32.573	0.706	0.407
Catch	1	31.991	0.048	0.828
CPUE	1	32.560	0.111	0.741

Note. An asterisk (*) indicates a significant value (alpha = 0.05).

4.2.1.4 Comparing the proportion of anglers in the first contact and follow-up contact within group

Chi-square goodness of fit tests provided evidence of a difference between the proportion of anglers, by type of angler, in the first contact and follow-up survey for Gp1_electronic (X^2 (2) = 29.6, p < 0.0001) and Gp2_sample (X^2 (2) = 7.2, p = 0.03), respectively. The chi-square goodness of fit test did not provide evidence of a difference for the Gp1_paper survey (X^2 (2) = 4.4, p = 0.12) although the mosaic plot (see Appendix F) indicated that the proportions were likely different. The result provided by the chi-square goodness of fit test for the Gp1_paper survey was attributed to the fact that the “did-not-fish” group only had 5 members. The test likely lacked statistical power.

4.2.1.5 Comparing fishing estimates and sample representativeness by survey mode: regular mail, email, and telephone

Welch’s test on ranked transformed data provided evidence that at least one survey mode consisted estimates that were from a different distribution for: total effort (Welch’s F (2, 227.4) = 5.64, p = 0.004), total catch (Welch’s F (2, 224.6) = 9.52, p < 0.0001), and CPUE (Welch’s F (2, 227.2) = 5.63, p = 0.004). This difference was identified between the email and the regular mail survey for: total effort (p = 0.003, median email = 4 angler-days, median regular mail = 3 angler-days), total catch (p < 0.0001, median email = 4 white sturgeon, median regular mail = 2 white sturgeon), and CPUE (p = 0.006, median email = 1 white sturgeon per angler-day, median regular mail = 0.7 white sturgeon per angler-day). Dunn’s multiple comparison post-hoc tests did not provide evidence of a clear difference between follow-up survey and the email or regular mail survey, in each case. Data were grouped by survey mode (i.e., regular mail, email, and telephone) for this analysis.

The proportions of anglers, by residency, in Gp1_electronic were different from Group_1 population (X^2 (2) = 22.4, p < 0.0001) (see Appendix G). Non-Canadian anglers were less likely to respond to the electronic survey. Chi-square goodness of fit test did not provide evidence for Gp1_paper (X^2 (2) = 1.6, p = 0.45) or Gp2_paper (X^2 (2) = 2, p = 0.37).

4.3 Sensitivity analysis

Three predominant trends were observed in the resampling exercise (see Appendix H for annual total effort and annual total catch plots). First, results indicated that the follow-up survey was important in reducing the likelihood of underestimating CPUE (Figure 3). In this, CPUE was underestimated when resampling was done with full samples from first contacts while excluding data from follow-up survey (Figure 3, upper right pane). Second, estimates behaved more stable when the sample size of the follow-up survey was held constant (Figure 3, middle pane). This behaviour held true even when sample size in the first contacts was reduced by 40%.

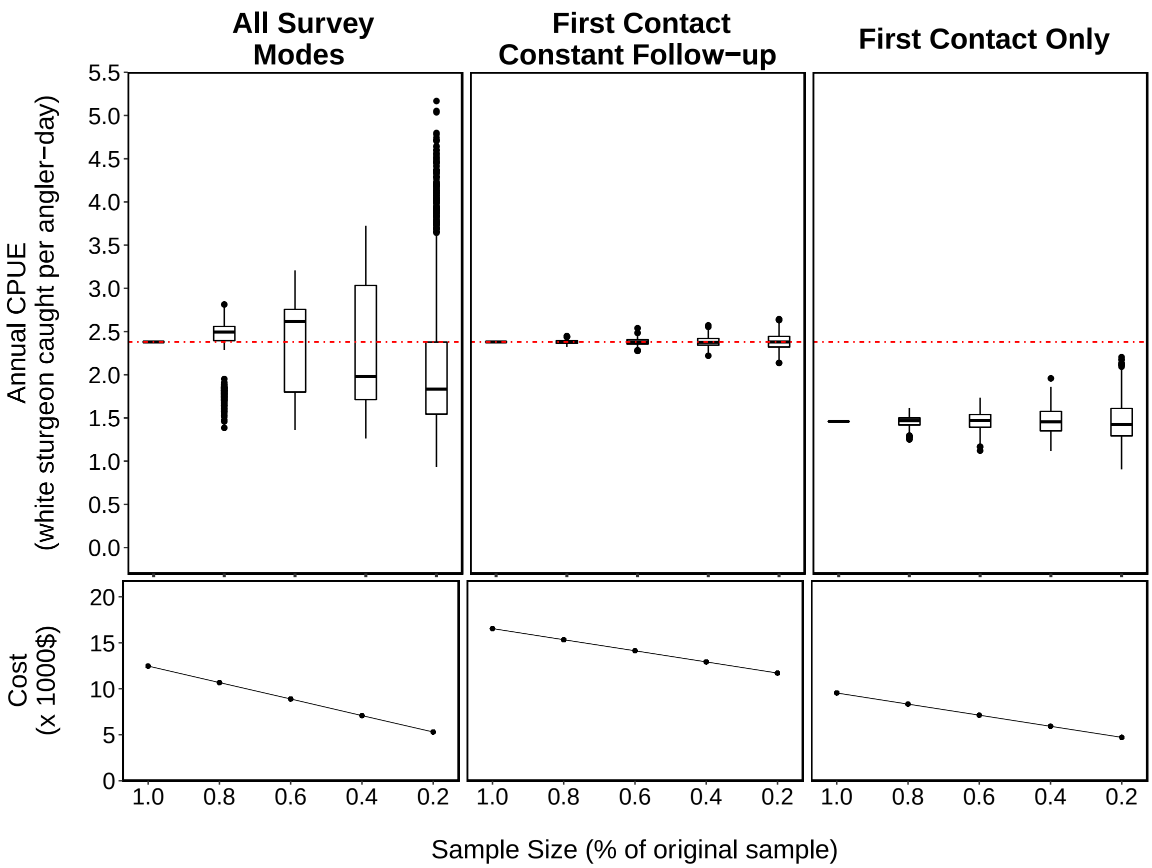


Figure 3. Results of the resampling exercise. Variation in CPUE and survey cost as a function of sample size and survey effort.

Note. The dotted horizontal line is a reference line. This reference line represents the estimated CPUE based on a resample of the complete data set.

Root mean square error for CPUE was lowest when the size of the follow-up survey was held constant (Figure 4). Third, survey cost and accuracy were not directly proportional.

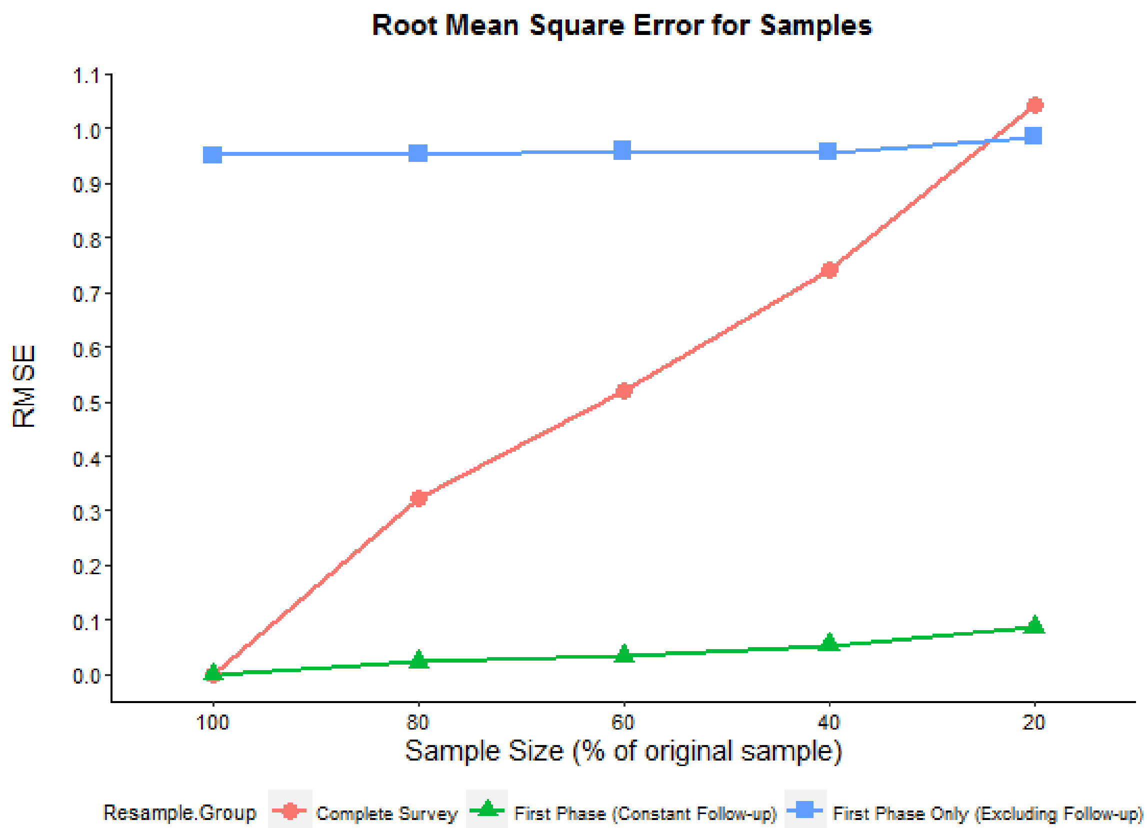


Figure 4. Root mean square error (RMSE) of samples' CPUE as a function of survey effort

5 Discussion

The purpose of this research was twofold. First, the research sought to estimate the effect of nonresponse on total effort and total catch estimates obtained from the sturgeon angler survey. Second, the research sought to quantify the accuracy and precision of estimates as a function of survey effort (i.e., sample size and the number of survey modes used to collect data). This section is presented in the following order: (1) I discuss the general findings; (2) I discuss the effects of nonresponse on survey estimates; (3) I discuss the effects of survey mode on survey estimates; and (4) I discuss the accuracy and precision of sample estimates as a function of survey effort.

In this research, 25% of non-guided anglers accounted for the majority of annual total catch and annual total effort. We classified anglers who were in this 25% group as “avid white surgeon anglers”. A few of these avid white sturgeon anglers reported extremely high individual total effort and total catch values that drew our attention. While it is common for avid anglers to have significantly higher catch and effort values than less avid anglers (Baccante, 1995), a portion of catch and effort values reported in the sturgeon angler survey seemed very unusual. Some of these effort and catch values might have affected by recall and/or prestige bias (Hartill & Edwards, 2015; Zarauz et al., 2015). Although I did not have evidence to support the presence of these biases, I theorized that sturgeon angler survey estimates are likely affected by recall bias because of the relatively long recall period (i.e., up to thirteen months) and prestige bias because of the likelihood of anglers over reporting catch to appear successful. Despite the uncertainties surrounding the plausibility of these extreme values, I did not remove them from my calculations because they could not be falsified. Furthermore, if estimates were truly affected by recall and prestige bias, an argument could have been made that all estimates (not only extreme values) were affected by biases. In essence, I used all values that were reported on completed surveys.

I observed nonresponse in the composition of anglers (i.e., guided anglers, non-guided anglers, and anglers who did not fish), anglers' effort, and to a very small extent in anglers' catch estimates. Only one survey (i.e., Gp1_paper) provided statistically different catch estimates between the first contact and the follow-up contact. On the other hand, results of the chi-square test provided evidence that the composition of anglers, by type of angler (i.e., guided, non-guided, and anglers who did not fish), was statistically different between first contact and follow-up contact for Gp1_electronic and Gp2_sample. The proportion of anglers who did not fish was higher (11 percentage points higher in Gp1_electronic and 8% percentage points higher in Gp2_paper) in the follow-up survey than in the first contact. This indicated that anglers who did not fish were less likely to respond to the fist contact surveys. If I extrapolated survey estimates to the population based on the first contact alone I would have obtained incorrect estimates. This finding is in line with Fisher (1996) and Zarauz el at. (2015) who reported that respondents had a higher participation rate than nonrespondents. This finding solidifies the need to explore off-site angler surveys for nonresponse bias, especially when response rate is low (Fisher, 1996).

A low response rate violates the central idea of selecting a random sample since respondents are somewhat self-selected into the sample (Brick, 2013). Self-selection affects the representativeness of samples (Petrov Ci, Petri, & Manfreda, 2016). For example, avid anglers are often overrepresented in fisheries surveys because they are more likely to respond (Thomson, 1991). I used anglers' licence type and residency as proxy variables for sample representativeness. Gp1_paper and Gp2_paper samples were representative of their respective population. However, Gp1_electronic had an underrepresentation of non-Canadian anglers and 1-Day licence holders in the responding sample; this is an indication that Gp1_electronic's responding sample was not representative of Gp1 population. It must be noted that the underrepresentation of non-Canadian anglers and 1-Day licence holders is only important if non-Canadian anglers and 1-Day licence holders had unique fishing patterns compared to anglers with other residency type and licence type. For example, if 1_Day licence holders tended to be less avid anglers who caught less fish than 8-day licence holders and/or annual licence holders, our estimates would be overestimated. Given this potential adverse effect of sample representativeness, I explored the dataset to determine the cause of the underrepresentation of Canadian anglers and 1-Day licence holders in Gp1_electronic's responding sample.

I found that coverage error played a role in the poor sample representativeness. The follow-up survey indicated that 63% of Gp1_electronic follow-up survey respondents did not receive a questionnaire in the first phase of contact. I did not have resources to explore the reasons for the potential low contact rate. However, in general, potential causes include using invalid, incorrect, or old email addresses, technical problems with the communication path of the emails, and personal and organizational email settings that identified emails as spam (Callegaro, Manfreda, & Vehovar, 2015). If these reasons hold true, then it can be argued that respondents were not completely self-selected but were not contacted and thus not given the chance to respond. It might be important to inform anglers about the importance of providing a correct and up to date email address when they are purchasing their licence.

Survey mode tended to have an effect on survey estimates and response rates. While the literature identifies the benefits to be gained from mixed mode surveys (see Dillman et al., 2014; Gigliotti & Dietsch, 2014; Vaske et al., 2011) as used here, Vannieuwenhuyze, Loosveldt, & Molenberghs (2010) pointed out that a single participant was likely to give varied responses to the same question when contacted via different survey modes; this phenomenon of obtaining a varied response is referred to as a mode effect. Evidence of this mode effect is abound in the literature and was explored in Chapter 2. For example, Zarauz et al. (2015) reported that a regular mail survey, email survey, and telephone survey yielded 351, 129, and 156 metric tonnes of catch, respectively, for the same sample. Laborde et al. (2014) also reported that respondents to an internet survey hunted more often, harvested more waterfowls, and placed greater importance on waterfowl hunting than respondents to a similar mail survey. Based on these findings in the literature, I hypothesized that at least one survey would produce statistically different estimates from the other two modes. Statistical test provided evidence of a difference between the regular mail survey and the email survey. Similar to Laborde et al. (2014), estimates from the email survey conducted in the present study tended to be higher than estimates from the regular mail survey. Presence of this mode effect is substantiated by the fact that estimates obtained for Gp1_paper tended to be closer to Gp2_paper estimates and further from Gp1_electronic estimates. Gp1_paper and Gp1_electronic were selected from the same population. However, similar to Gp2_paper, members of Gp1_paper received a questionnaire via regular mail. An inference can be made that survey mode had an effect on respondents and the estimates that they reported rather than anglers from Gp1 and Gp2 having different fishing behavior.

Within this research, the regular mail survey outperformed the email survey on 3 of 4 criteria (Table 4). I found that the email survey was cheaper but was outperformed by the mail and telephone survey in terms of response rate and sample representativeness. This finding is similar to Zarauz et al. (2015). An important aspect of the email survey is that the entire sample can be contacted for the same upfront cost. Therefore, a large sample can be contacted and the data can be entered with little to no additional cost. However, the low response rate and significant bias estimates might be a deterrent.

Table 4. Comparison of email and regular mail survey based on 4 criteria

	Survey Mode	
Criteria	**Electronic**	**Regular Mail**
Cost	✓	
Representative Sample		✓
Response Rate		✓
Successful Contact		✓

The resampling exercise indicated that the variance of estimates and sample size were inversely related. The large variation of estimates is attributed to the large variance in data reported by anglers. As seen in Figure 3, resampled estimates behaved less erratic when the sample size for follow-up survey was held constant. This behaviour held true even when sample size was reduced to 20%. The RMSE, being an indicator for the accuracy of estimates (Walther & Moore, 2005), indicated that the estimates were most accurate when the sample size of the follow-up survey was held constant. Since the Welch's test did not provide evidence of a difference between the first and follow-up contact, it was expected that the sample size could be reduced without any great effect on accuracy; however, this was not the case. Most of the risk associated with reducing the sample size can be attribute with the likelihood of not obtaining a representative sample. Here, the criteria for a representative sample include obtaining estimates with a similar distribution and actually capturing non-guided anglers in the responding sample.

5.1 Limitations

A major limitation of this research was the unavailability of data independent of the survey (for example data from a creel survey) to verify catch and effort estimates reported by anglers. Independent data could have been used to verify if extreme catch and effort values were plausible. Another limitation was manifested in the absence of auxiliary data. As pointed out by Lew et al. (2015) and Lohr (2010), auxiliary data can be used in survey poststratification to reduce variance, improve the precision of estimates, and mitigate the effect of nonresponse. More specific to this research, auxiliary data could have been used to form weighting classes and reduce the effect of outlier values on the mean by assigning anglers with similar characteristics to the same weighting class. Of course, auxiliary variables used to create weighting classes would need to have a strong correlation with estimates. The final notable limitation arose from the fact that follow-up survey did not achieve a 100% response rate. There is therefore a group of nonrespondents from who no data were obtained.

6 Conclusion

In conclusion, the findings of this present study reinforce the need to explore the The Annual Fraser River White Sturgeon Angling Questionnaire for nonresponse bias. This finding can be extended to similar off-site recreational fisheries surveys. I found that nonresponse bias arose from anglers' participation rate and to a lesser extent in anglers catch estimates. I also found that the regular mail survey outperformed the email survey on three out of four criteria, namely: sample representativeness, contact rate, and response rate. My findings reinforced the superiority of regular mail surveys over email surveys by showing that this superiority holds true even in a recreational fishery survey of a relatively small group of anglers, and with a complete sample frame being available. Results from the resampling exercise reinforced the importance of the follow-up survey. Finally, I found that sample size in the first phase of contact can be reduced by 40 percentage points while holding the following survey at the original sample size and still yield accurate results.

References

American Fisheries Society. (2011). New Frontiers in Fisheries Management and Ecology: Leading the Way in Changing the World. American Fisheries Society 141st Annual Meeting. Retrieved from https://afs.confex.com/afs/2011/webprogram/authora.html

Assael, H., & Keon, J. (1982). Nonsampling vs. Sampling Errors in Survey Research. *Journal of Marketing*, *46*(2), 114–123. http://doi.org/10.2307/3203346

Baccante, D. (1995). Assessing catch inequality in walleye angling fisheries. *North American Journal of Fisheries Management*, *15*(3), 661 –665. http://doi.org/0.1577/1548-8675(1995)015<0661:ACIIWA>2.3.CO;2

Bai, H., & Pan, W. (2008). Resampling methods revisited: advancing the understanding and applications in educational research. *International Journal of Research & Method in Education*, *31*(1), 45–62. http://doi.org/10.1080/17437270801919909

Barnes, M. E., Simpson, G., Carreiro, J., & Voorhees, J. (2014). A Comparison of A Creel Census to Modeled Access-Point Creel Surveys on Two Small Lakes Managed as Put-and-Take Rainbow Trout Fisheries. *Fisheries and Aquaculture Journal*, *05*(01). http://doi.org/10.4172/2150-3508.1000086

Beaman, J., Vaske, J. J., Schmidt, J. I., & Huan, T. (2015). Measuring and Correcting Response Heaping Arising From the Use of Prototypes. *Human Dimensions of Wildlife*, *20*(2), 167–173. http://doi.org/10.1080/10871209.2015.968890

Beardmore, B., Hunt, L. M., Haider, W., Dorow, M., & Arlinghaus, R. (2014). Effectively managing angler satisfaction in recreational fisheries requires understanding the fish species and the anglers. *Canadian Journal of Fisheries and Aquatic Sciences*. Retrieved from http://www.nrcresearchpress.com/doi/full/10.1139/cjfas-2014-0177#.VstJvlsrJaQ

Brick, J. M. (2013). Unit nonresponse and weighting adjustments: A critical review. *Journal of Official Statistics*, *29*(3), 329–353. http://doi.org/10.2478/jos-2013-0026

Brick, J. M., & Williams, D. (2012). Explaining Rising Nonresponse Rates in Cross-Sectional Surveys. *The ANNALS of the American Academy of Political and Social Science*, *645*(1), 36–59. http://doi.org/10.1177/0002716212456834

Callegaro, M., Manfreda, K. L., & Vehovar, V. (2015). *Web Survey Methodology*. SAGE Publications. Retrieved from https://books.google.com/books?hl=en&lr=&id=RvsaCAAAQBAJ&pgis=1

Carkin, D. M., & Tracy, P. E. (2015). Adjusting for Unit Non-Response in Surveys Through Weighting. *Crime & Delinquency*, *61*(1), 143–158. http://doi.org/10.1177/0011128714556739

Committee on the Status of Endangered Wildlife in Canada. (2012). *COSEWIC assessment and status report on the White Sturgeon Acipenser transmontanus in Canada*. Retrieved from http://publications.gc.ca/collections/collection_2013/ec/CW69-14-353-2013-eng.pdf

Connelly, N. A., & Brown, T. L. (2011). Effect of recall period on annual freshwater fishing effort estimates in New York. *Fisheries Management and Ecology*, *18*(1), 83–87. http://doi.org/10.1111/j.1365-2400.2010.00777.x

Connelly, N. A., Brown, T. L., & Knuth, B. A. (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. http://doi.org/10.1080/10871200009359192

Dillman, D. A. (2011). *Mail and Internet surveys: The tailored design method--2007 Update with new Internet, visual, and mixed-mode guide*. John Wiley & Sons.

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: the tailored design method.* John Wiley & Sons.

English, K. K., & Jesson, D. (2013). *Sturgeon Catch and Effort Estimates for Lower & Middle Fraser River Recreational Fishery 2009-12*.

Fedler, A. J., & Ditton, R. B. (1994). Understanding Angler Motivations in Fisheries Management. *Fisheries*, *19*(4), 6–13. http://doi.org/10.1577/1548-8446(1994)019<0006:UAMIFM>2.0.CO;2

Fisher, M. R. (1996). Estimating the Effect of Nonresponse Bias on Angler Surveys. *Transactions of the American Fisheries Society*, *125*(1), 118–126. http://doi.org/10.1577/1548-8659(1996)125<0118:ETEONB>2.3.CO;2

Fowler Jr, F. J. (2013). *Survey research methods*. Sage publications.

Fuchs, M., Bossert, D., & Stukowski, S. (2013). Response Rate and Nonresponse Bias - Impact of the Number of Contact Attempts on Data Quality in the European Social Survey. *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique*, *117*(1), 26–45. http://doi.org/10.1177/0759106312465547

Gigliotti, L., & Dietsch, A. (2014). Does Age Matter? The Influence of Age on Response Rates in a Mixed-Mode Survey. *Human Dimensions of Wildlife*, *19*(3), 280–287. http://doi.org/10.1080/10871209.2014.880137

Groves, R. M. (1989). *Survey errors and survey costs. Wiley Series in Probability and Mathematical Statistics*. New York, NY: Wiley.

Groves, R. M., & Heeringa, S. G. (2006). Responsive design for household surveys: tools for actively controlling survey errors and costs. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *169*(3), 439–457. http://doi.org/10.1111/j.1467-985X.2006.00423.x

Groves, R. M., Presser, S., & Dipko, S. (2004). the Role of Topic Interest in Survey Participation Decisions. *The Public Opinion Quarterly*, *68*(1), 2–31. http://doi.org/doi:10.1093/poq/nfh002

Groves, R. M., Singer, E., & Corning, A. (2000). Leverage-saliency theory of survey participation: description and an illustration. *He Public Opinion Quarterly*, *64*(3), 299– 308. Retrieved from http://www.jstor.org.proxy.lib.sfu.ca/stable/pdf/3078721.pdf?acceptTC=true

Hansen, M. . ., & Hurwitz, W. . . (1946). The Problem of Non-Response in Sample Surveys. *Journal of the American Statistical Association*, *41*(236), 517–529. http://doi.org/10.2307/2280572

Hartill, B. W., & Edwards, C. T. T. (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*(9), 1379–1389. http://doi.org/10.1139/cjfas-2014-0451

Hoyle, S. D., & Cameron, D. S. (2003). Confidence intervals on catch estimates from a recreationsl fishing survey: a comparison of bootstrap methods. *Fisheries Management and Ecology*, *10*, 97–108. http://doi.org/10.1046/j.1365-2400.2003.00321.x

Johnston, F. D., Arlinghaus, R., & Dieckmann, U. (2010). Diversity and complexity of angler behaviour drive socially optimal input and output regulations in a bioeconomic recreational-fisheries model. *Canadian Journal of Fisheries and Aquatic Sciences*, *67*(9), 1507–1531. http://doi.org/10.1139/F10-046

Kent, R. (2001). *Data construction and data analysis for survey research*. Macmillan.

Krueger, B. S., & West, B. T. (2014). Assessing the Potential of Paradata and Other Auxiliary Data for Nonresponse Adjustments. *Public Opinion Quarterly*, *78*(4), 795–831. http://doi.org/10.1093/poq/nfu040

Laborde, L. P., Rohwer, F. C., Kaller, M. D., & Reynolds, L. A. (2014). Surveying Louisiana waterfowl hunters: Open web and random mail surveys produce similar responses to attitudinal questions. *Wildlife Society Bulletin*, *38*(4), 821–826. http://doi.org/10.1002/wsb.492

Larkin, M. F., Ault, J. S., Humston, R., & Lou, L. (2010). A mail survey to estimate the fishery dynamics of southern Florida's bonefish charter fleet. *Fisheries Management and Ecology*, *17*(3), 254–261. http://doi.org/10.1111/j.1365-2400.2009.00718.x

Lew, D. K., Himes-Cornell, A., & Lee, J. (2015). Weighting and imputation for missing data in a cost and earnings fishery survey. *Marine Resource Economics*, *30*(2), 219–230. http://doi.org/10.1086/679975

Littvay, L., Popa, S. A., & Fazekas, Z. (2013). Validity of Survey Response Propensity Indicators: A Behavior Genetics Approach. *Social Science Quarterly*, *94*(2), 569–589. http://doi.org/10.1111/j.1540-6237.2012.00938.x

Lohr, S. L. (2010). *Sampling: Design and analysis (2nd. ed.).* Boston, MA: Brooks/Cole.

Lowry, D., Stick, K., Lindquist, A., & Cheng, Y. W. (2015). Evaluation of Creel Survey Methods to Estimate Recreational Harvest of Surf Smelt in Puget Sound, Washington. *North American Journal of Fisheries Management*, *35*(3), 403–417. http://doi.org/10.1080/02755947.2015.1009658

Marsden, P. V., & Wright, J. D. (2010). *Handbook of survey research*. Emerald Group Publishing.

McCormick, J. L., Quist, M. C., & Schill, D. J. (2013). Self-Reporting Bias in Chinook Salmon Sport Fisheries in Idaho: Implications for Roving Creel Surveys. *North American Journal of Fisheries Management*, *33*(4), 723–731. http://doi.org/10.1080/02755947.2013.808293

Mccormick, J. L., Whitney, D., Schill, D. J., & Quist, M. C. (2015). Evaluation of angler reporting accuracy in an off-site survey to estimate statewide steelhead harvest. *Fisheries Management and Ecology*, *22*(2), 134–142. http://doi.org/10.1111/fme.12109

McNabb, D. E. (2013). *Nonsampling error in social surveys*. Sage publications.

Ministry of Forests Lands and Natural Resource Operations. (2015). Angling Licences and Fees. White Sturgeon Conservation Licence. Retrieved from http://www.jstor.org/stable/3078721

Mooney, C. Z., Duval, R. D., & Duval, R. (1993). *Bootstrapping: A nonparametric approach to statistical inference (No. 94-95)*. Sage.

National Research Council. (2006). *Committee on the Review of Recreational Fisheries Survey Methods*. Retrieved from http://www.nmfs.noaa.gov/ocs/mafac/meetings/2006_07/docs/NRC_MRFSS_2006.pdf

Needham, M. D., Scott, D., & Vaske, J. (2013). Recreation Specialization and Related Concepts in Leisure Research. *Leisure Sciences*, *35*(3), 199–202. http://doi.org/10.1080/01490400.2013.780457

Peress, M. (2010). Correcting for Survey Nonresponse Using Variable Response Propensity. *Journal of the American Statistical Association*, *105*(492), 1418–1430. http://doi.org/10.1198/jasa.2010.ap09485

Petrov Ci, A. Z., Petri, G., & Manfreda, K. L. (2016). The effect of email invitation elements on response rate in a web survey within an online community. *Computers in Human Behavior*, *56*, 320–329. http://doi.org/10.1016/j.chb.2015.11.025

Peytchev, A. (2013). Consequences of survey nonresponse. *The ANNALS of the American Academy of Political and Social Science*, *645*(1), 88–111. http://doi.org/10.1177/0002716212461748

Pollock, K. H., Jones, C. M., & Brown, T. L. (1994). *Angler survey methods and their application in fisheries management*. American Fisheries Society Special Publication 25.

Puth, M. T., Neuhauser, M., & Ruxton, G. D. (2015). On the variety of methods for calculating confidence intervals by bootstrapping. *Journal of Animal Ecology*, *84*(4), 892–897. http://doi.org/10.1111/1365-2656.12382

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*(4), 688–690. http://doi.org/10.1093/beheco/ark016

Singer, E., & Ye, C. (2012). The Use and Effects of Incentives in Surveys. *The ANNALS of the American Academy of Political and Social Science*, *645*(1), 112–141. http://doi.org/10.1177/0002716212458082

Sudman, S., Bradburn, N. M., & Schwarz, N. (1996). *Thinking about answers: The application of cognitive processes to survey methodology*. Jossey-Bass.

Sullivan, M. G. (2003). Exaggeration of Walleye Catches by Alberta Anglers. *North American Journal of Fisheries Management*, *23*(2), 573–580. http://doi.org/10.1577/1548-8675(2003)023<0573:EOWCBA>2.0.CO;2

Tarrant, M. A., Manfredo, M. J., Bayley, P. B., & Hess, R. (1993). Effects of Recall Bias and Nonresponse Bias on Self-Report Estimates of Angling Participation. *North American Journal of Fisheries Management*, *13*(2), 217–222. http://doi.org/10.1577/1548-8675(1993)013<0217:EORBAN>2.3.CO;2

Thompson, S. K. (2012). *Sampling (3rd ed.)*. Hoboken, N.J: Wiley.

Thomson, C. J. (1991). Effects of the avidity bias on survey estimates of fishing effort and economic value. *American Fisheries Society Symposium*, *12*(356366).

Valliant, R., Dever, J. A., & Kreuter, F. (2013). *Practical tools for designing and weighting survey samples*. New York: Springer.

Vannieuwenhuyze, J., Loosveldt, G., & Molenberghs, G. (2010). A method for evaluating mode effects in mixed-mode surveys. *Public Opinion Quarterly*, *74*(5), 1027–1045. http://doi.org/10.1093/poq/nfq059

Vaske, J. J., & Beaman, J. (2006). Lessons Learned in Detecting and Correcting Response Heaping: Conceptual, Methodological, and Empirical Observations. *Human Dimensions of Wildlife*, *11*(4), 285–296. http://doi.org/10.1080/10871200600803234

Vaske, J. J., Huan, T., & Beaman, J. (2003). The Use of Multiples in Anglers' Recall of Participation and Harvest Estimates: Some Results and Implications. *Leisure Sciences*, *25*(4), 399–409. http://doi.org/10.1080/714044498

Vaske, J. J., Jacobs, M. H., Sijtsma, M. T. J., & Beaman, J. (2011). Can Weighting Compensate for Sampling Issues in Internet Surveys? *Human Dimensions of Wildlife*, *16*(May 2015), 200–215. http://doi.org/10.1080/10871209.2011.571230

Walther, B. A., & Moore, J. L. (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*(6), 815–829. http://doi.org/10.1111/j.2005.0906-7590.04112.x

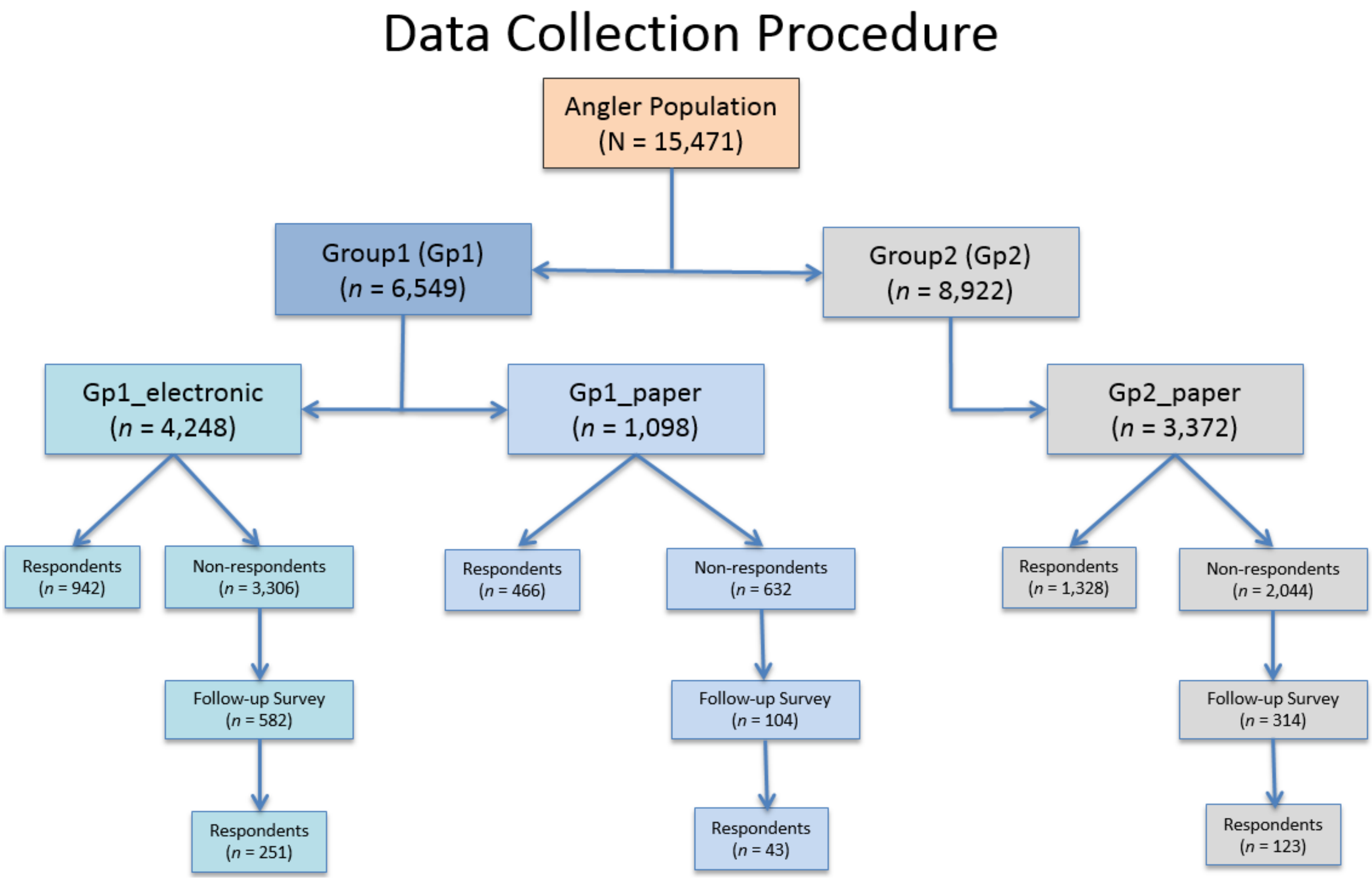
Xu, B., Zhang, C., Xue, Y., Ren, Y., & Chen, Y. (2015). Optimization of sampling effort for a fishery-independent survey with multiple goals. *Environmental Monitoring and Assessment*, *187*(5), 252. http://doi.org/10.1007/s10661-015-4483-9

Zarauz, L., Ruiz, J., Urtizberea, A., Andonegi, E., Mugerza, E., & Artetxe, I. (2015). Comparing different survey methods to estimate European sea bass recreational catches in the Basque Country. *ICES Journal of Marine Science*, *72*(4), 1181–1191. http://doi.org/10.1093/icesjms/fsv054

Zimmerman, D. W. (2012). A note on consistency of non-parametric rank tests and related rank transformations. *The British Journal of Mathematical and Statistical Psychology*, *65*(1), 122–44. http://doi.org/10.1111/j.2044-8317.2011.02017.x

Appendix A.

Schematic Diagram of Data Collection Procedure



The diagram above illustrates the data collection procedure used in this project. The sample frame was the White Sturgeon Conservation Licence database for the 2014/15 white sturgeon fishing season. The White Sturgeon Conservation Licence database contained the name and contact information of all anglers who purchased a White Sturgeon Conservation Licence. Data collection procedure within each group followed a two-phase sample design. The data collection procedure followed the following steps:

1. The population was divided into two mutually exclusive groups, herein called Group1 (Gp1) and Group2 (Gp2). Gp1 consisted of anglers who, at the time of purchasing their WSCS, provided an email address and consented to being contacted via email, Gp2 consisted of anglers who did not provide an email address.

2. Two simple random samples were concurrently selected from Gp1; herein called Gp1_electronic and Gp1_paper. Members of Gp1_electronic were sent a questionnaire via email. Members of Gp1_paper were sent a questionnaire via regular mail.

3. A simple random sample was selected from Gp2 (i.e., Gp2_paper). Members of Gp2_paper were sent a questionnaire via regular mail.

4. Anglers who did not respond to the first phase were deemed nonrespondents. A simple random sample was selected from the nonrespondents in each group. The sample of nonrespondents was contacted via telephone. The follow-up contact was called the nonresponse assessment telephone survey.

Appendix B.

Annual Total Effort and Annual Total Catch Equations

Equation	Description	Eq. #
$\widehat{A_{total}} = \sum_{j=1}^{3} \hat{T}_j,$	where $\widehat{A_{total}}$ was the annual total estimate for the entire population and $\hat{T}_j$ was the annual total estimate for survey *j*.	1
$\hat{T} = \hat{t}_{first} + \hat{t}_{follow-up},$	where the annual total estimate for each survey ($\hat{T}$) was the sum of the annual total estimate reported in the first contact ($\hat{t}_{first}$) and the follow-up contact ($\hat{t}_{follow-up}$). As shown in Appendix A, each survey had two phases of contact.	2
$\hat{t}_i = N * W * \%Ng * \overline{est},$	where $\hat{t}_i$was the annual total estimate for each phase of contact, N was the group population, W was a weight adjustment, $\%Ng$ was the percentage of non-guided anglers in the responding sample, and $\overline{est}$ was the mean estimate per non-guided angler.	3
$W = \frac{q_C+q_I}{Q_s},$	where W was a weight adjustment that reflected the proportion of the sample that was represented by each phase of contact, q_C was the number of completed questionnaires returned, q_I was the number of incomplete questionnaires returned, and Q_s was the number of questionnaires sent (i.e., total sent – undeliverable). A completed questionnaire was classified as a questionnaire that had single, clearly legible response to all required questions. The respondents to the follow-up survey, for each survey, represented all non-respondents. Therefore, W for each follow-up survey was 1-$\frac{q_c+q_i}{Q_s}$.	4

$\%Ng = \frac{\mathrm{Ng}}{q_c}$,	where $\%Ng$ was the percentage of non-guided anglers in the sample, Ng was the number of non-guided anglers that responded, and q_C was the number of completed questionnaires that were returned.	5
$\overline{est} = \frac{1}{n} * \sum_{i=1}^{Ng} x_i$,	where $\overline{est}$ was the mean estimate per non-guided angler, Ng was the number of non-guided anglers, and x was the total catch or total effort reported by non-guided angler i.	6

Annual total estimate refers to annual total effort and annual total catch.

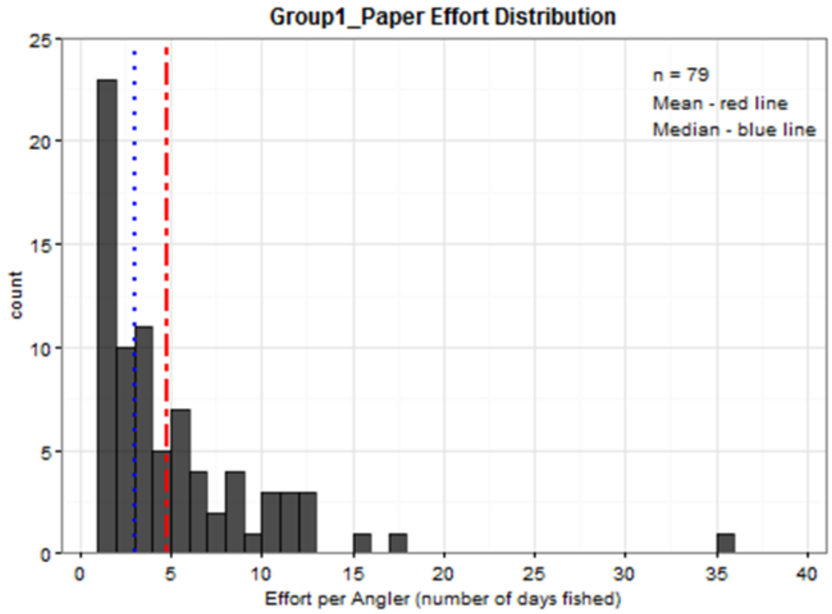
The table in Appendix B shows the equations used to calculate annual total estimates of effort and catch. The 2014-2015 sturgeon angler survey had three separate survey groups: Gp1_electronic, Gp1_paper, and Gp2_paper (see Appendix A). I calculated annual total estimates of effort and catch by adding the annual total estimates obtained from each survey group; this is shown in equation 1 in the table above and in the Methods section. Each survey group had two phases of contact, herein referred to as first contact and follow-up contact. I obtained the sample for each follow-up contact by selecting a simple random sample from the group of nonrespondents in each group. I made the assumption that data from each groups' follow-up contact was a true representation of all nonrespondents in that group. Therefore, the mean effort and mean catch from the follow-up survey were taken to be the mean for all non-respondents in that group. Equation 2 shows that the annual total estimate for each group was the sum of annual estimates obtained from the first contact and the follow-up contact.

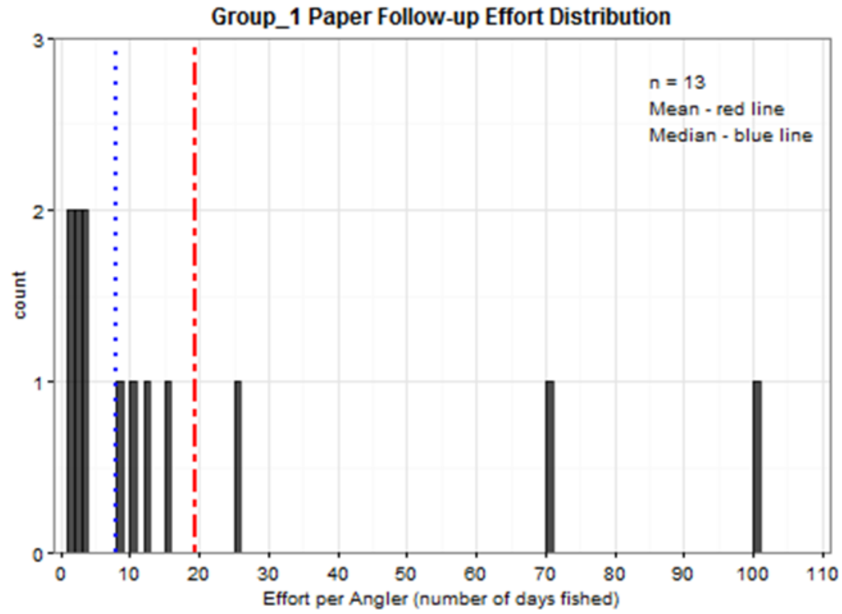
Equation 3 shows the equation used to calculate the annual total estimates for each phase within a group. The W is a weight adjustment. This adjustment gave a weight to the mean estimate from the first contact and a weight to the mean estimate from the follow-up contact when expanding the mean estimate to the subpopulation (see Appendix D). The W signifies what proportion of the sample was represented by data from the first contact (i.e., respondents) and what proportion of the sample was represented by data from the follow-up contact (i.e., non-respondents). To illustrate using and example, Gp2_paper sample size was 3,372. Forty-three percent of Gp2_paper sample returned a questionnaire in the first contact. The remaining 53% were nonrespondents. Therefore, within Gp2_paper, W was 0.43 for the first contact and 0.57 for the follow-up contact.

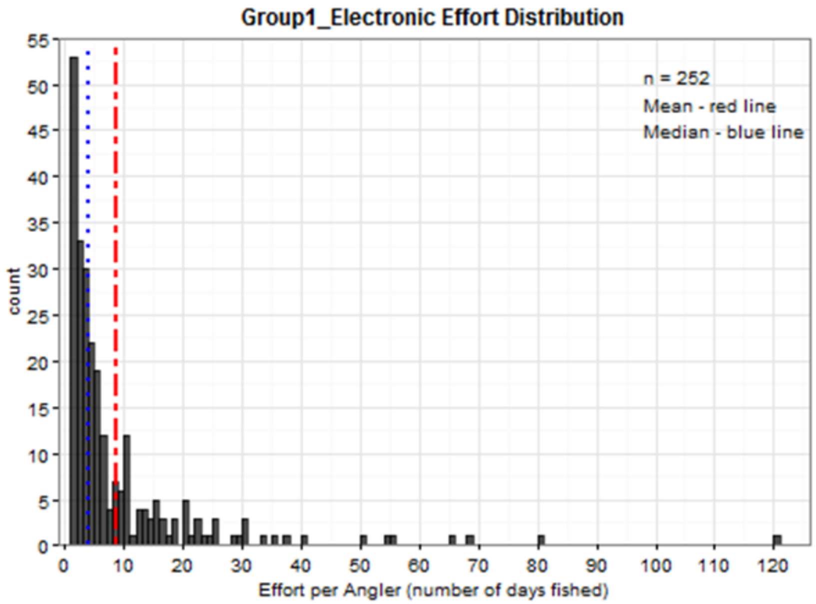
Appendix C.

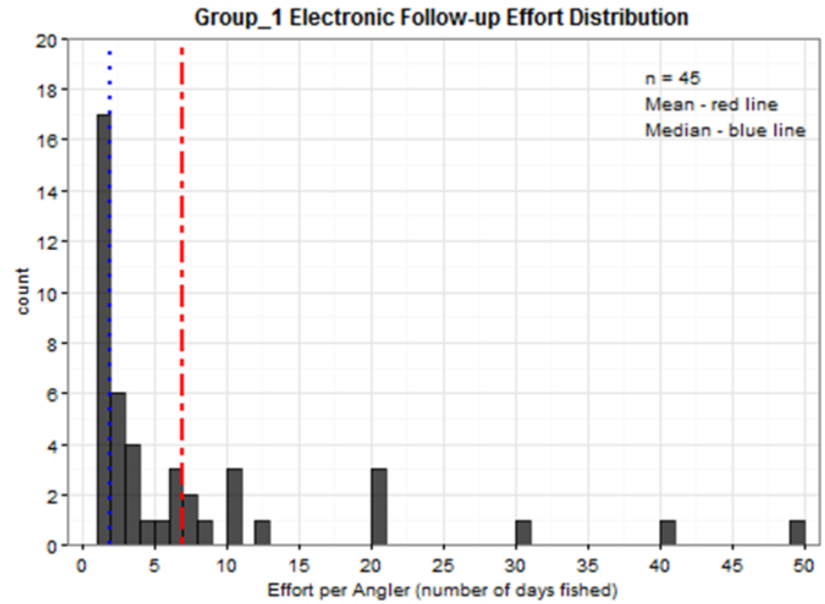
Catch and Effort Histograms

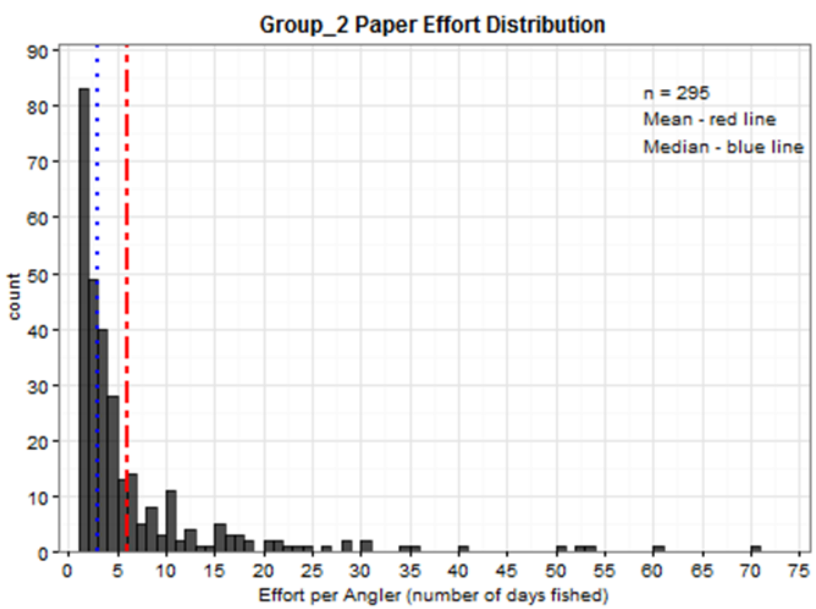
The histograms below present the distribution of data that were obtained from the surveys. Data were were highly right skewed and leptokurtic. All survey modes consisted of values that could be considered extreme values. However, these values were not removed because they could not be falsified.

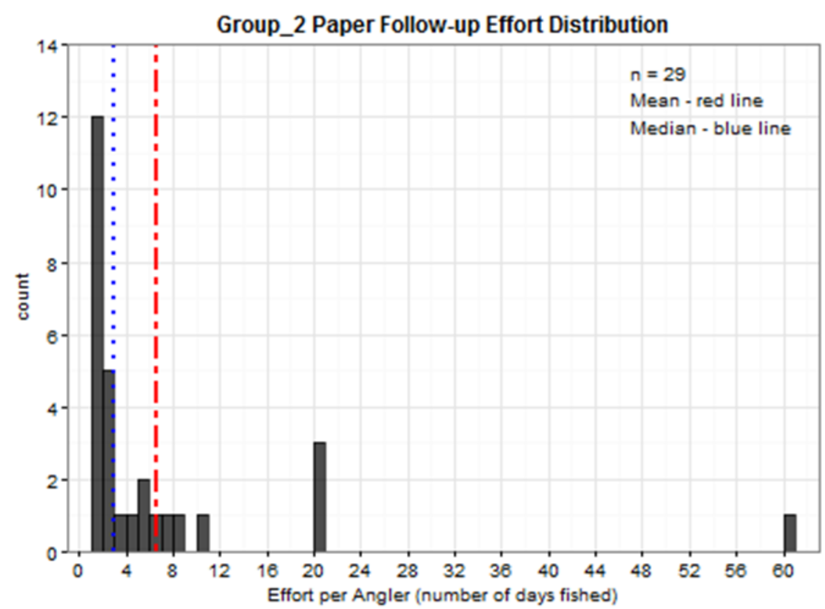


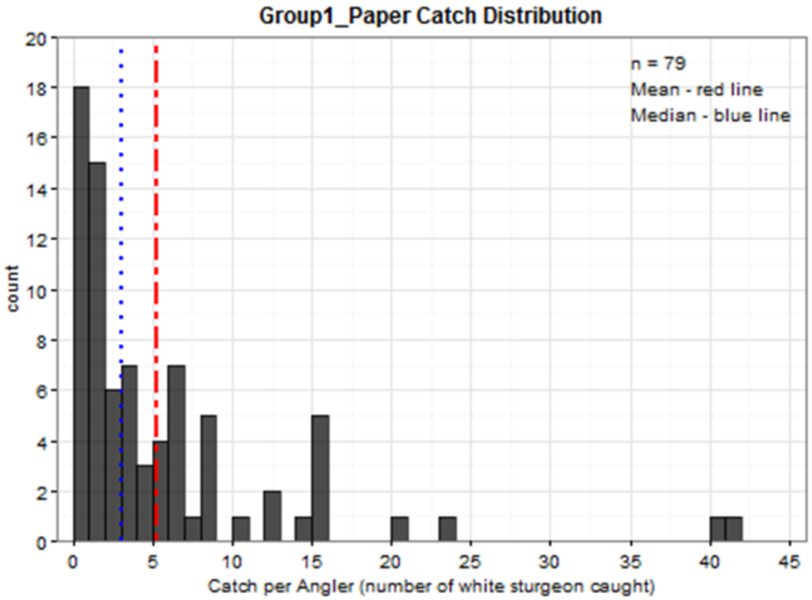


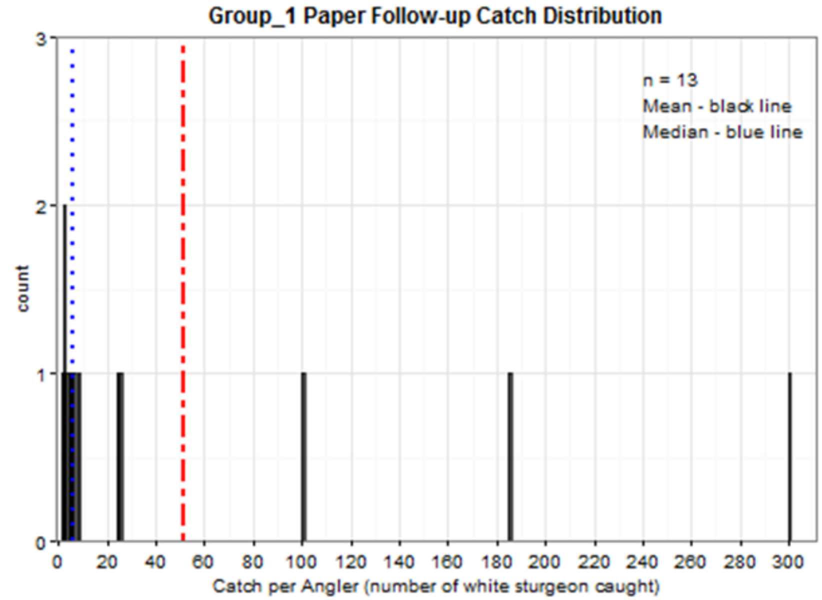

Group1_Electronic Effort Distribution
n = 252
Mean - red line
Median - blue line
count
Effort per Angler (number of days fished)

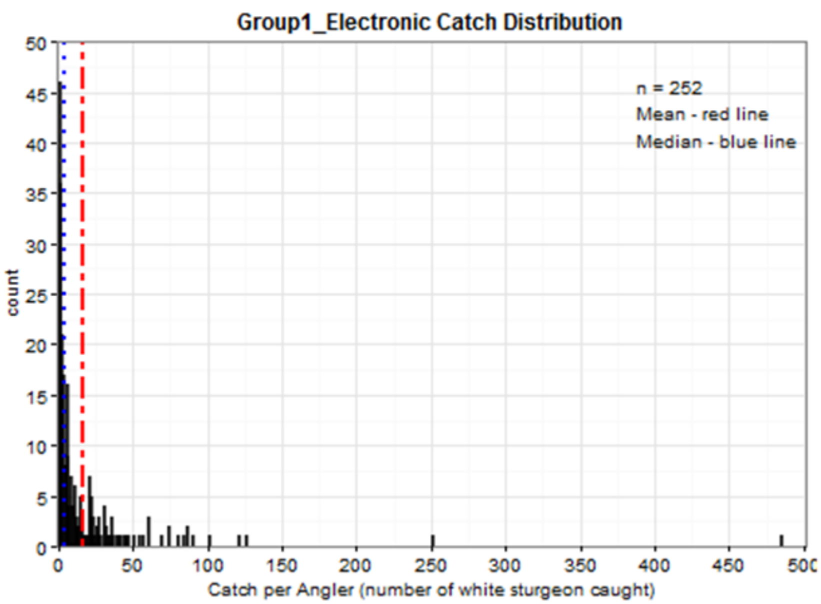

Group_1 Electronic Follow-up Effort Distribution
n = 45
Mean - red line
Median - blue line
count
Effort per Angler (number of days fished)

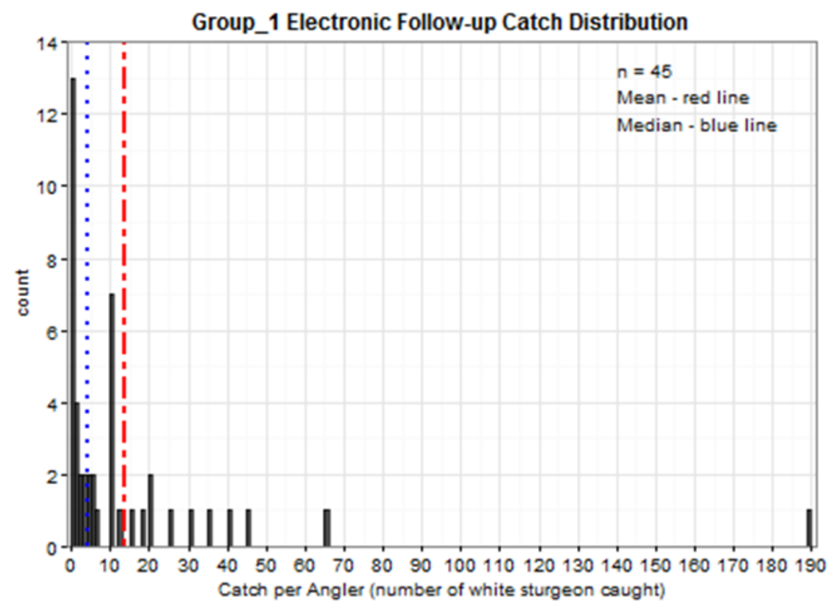

Group_2 Paper Effort Distribution
n = 295
Mean - red line
Median - blue line
count
Effort per Angler (number of days fished)

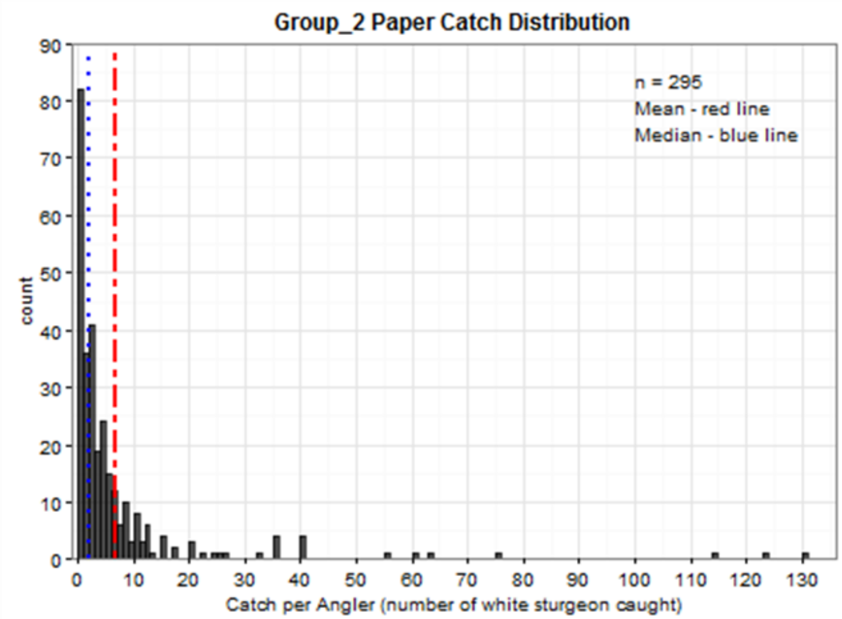

Group_2 Paper Follow-up Effort Distribution
n = 29
Mean - red line
Median - blue line
count
Effort per Angler (number of days fished)


Group1_Paper Catch Distribution
n = 79
Mean - red line
Median - blue line
count
Catch per Angler (number of white sturgeon caught)

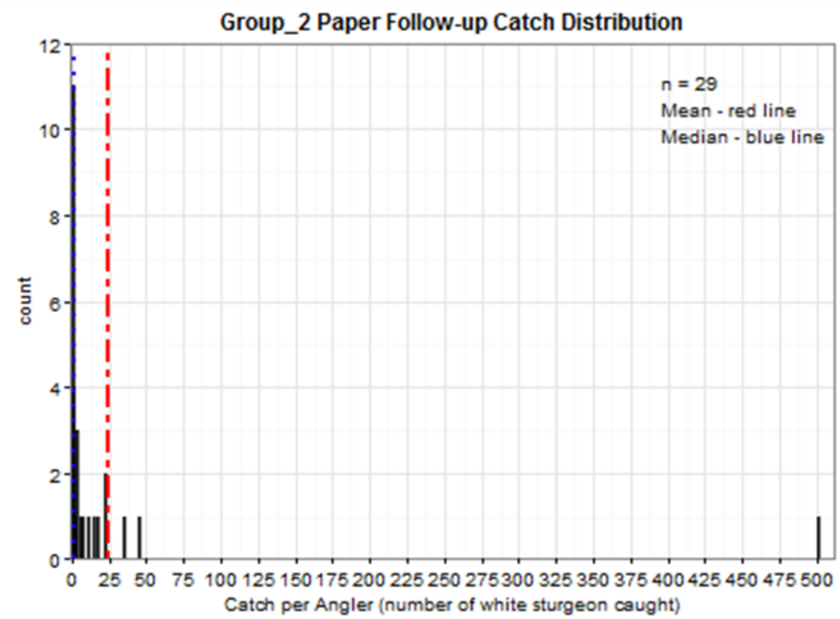

Group_1 Paper Follow-up Catch Distribution
n = 13
Mean - black line
Median - blue line
count
Catch per Angler (number of white sturgeon caught)


Group1_Electronic Catch Distribution
n = 252
Mean - red line
Median - blue line
count
Catch per Angler (number of white sturgeon caught)


Group_1 Electronic Follow-up Catch Distribution
n = 45
Mean - red line
Median - blue line
count
Catch per Angler (number of white sturgeon caught)



Group_2 Paper Catch Distribution
n = 295
Mean - red line
Median - blue line
count
0
10
20
30
40
50
60
70
80
90
0 10 20 30 40 50 60 70 80 90 100 110 120 130
Catch per Angler (number of white sturgeon caught)



Group_2 Paper Follow-up Catch Distribution
n = 29
Mean - red line
Median - blue line
count
0
2
4
6
8
10
12
0 25 50 75 100 125 150 175 200 225 250 275 300 325 350 375 400 425 450 475 500
Catch per Angler (number of white sturgeon caught)

Appendix D.

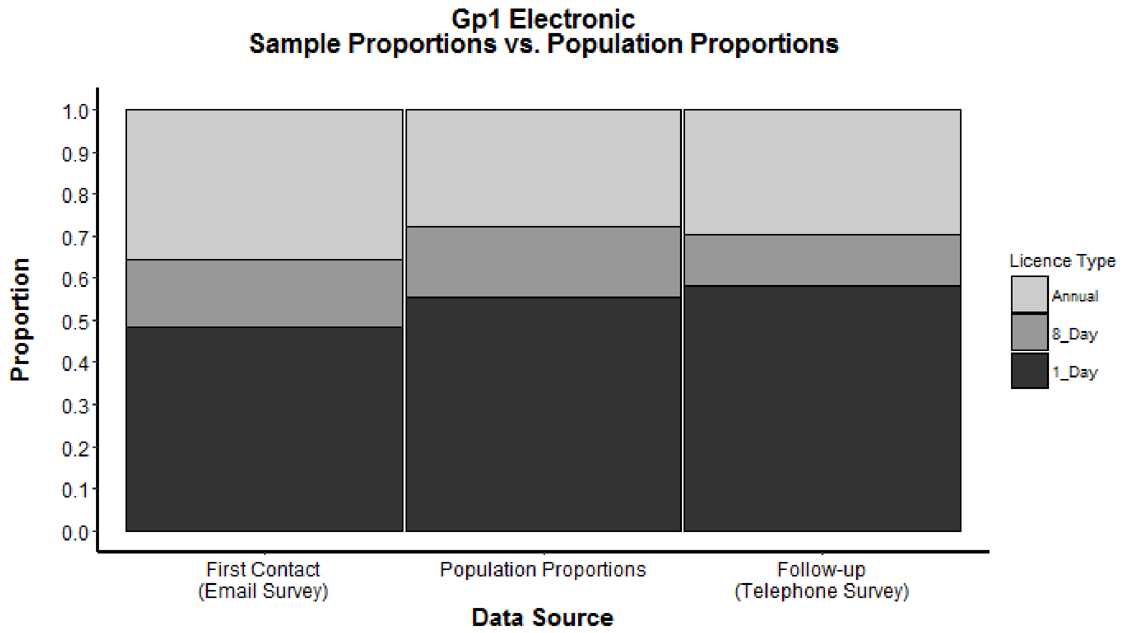
Annual total effort and annual total catch estimates and calculations

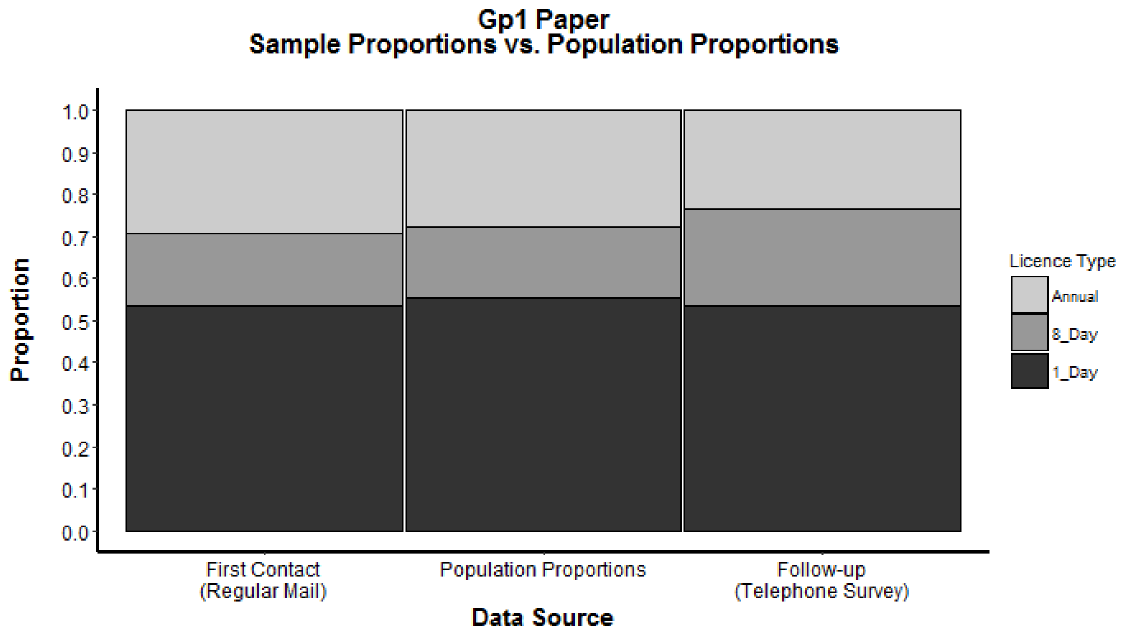
	Gp1_electronic		Gp1_paper		Gp2_paper	
Subpopulations (N)	5,204		1,345		8,922	
Sample size	4,248		1,098		3,372	
Return to sender	312		39		261	
Refusal/Spoil	89		9		49	
Actual contact	3,936		1,059		3,111	
Number of respondents (First Contact)	942		466		1,328	
	First contact	Follow-up	First contact	Follow-up	First contact	Follow-up
Weight (W)	0.24	0.76	0.44	0.56	0.43	0.57
Percentage of Non-guided anglers ($\%Ng$)	29%	18%	18%	30%	23%	23%
Mean effort ($\overline{est}$)	8.6	6.9	4.8	19.4	6.1	6.6
Mean catch ($\overline{est}$)	15.7	13.8	5.3	51.2	6.7	23.9
Annual total effort	3,106	4,916	511	4,383	5,343	7,762
Annual total catch	5,671	9,833	565	11,568	5,869	28,109
Annual Total Effort ($\widehat{A_{total}}$)	26,023	Angler-days				
Annual Total Catch ($\widehat{A_{total}}$)	61,614	White sturgeon				
CPUE	2.4					

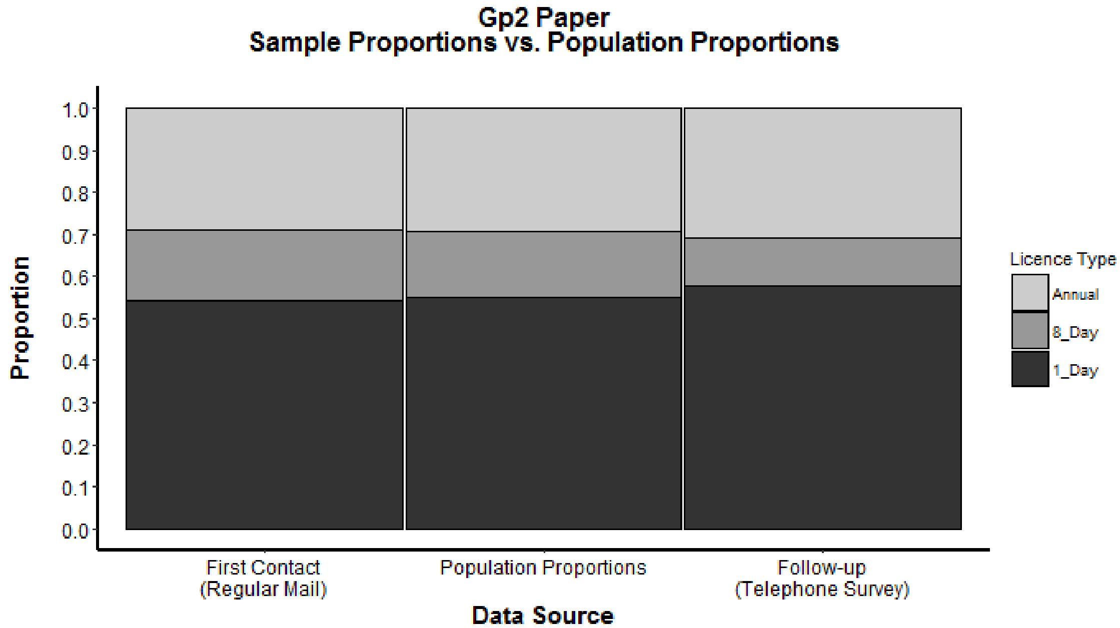
Estimated annual total effort and annual total catch are shown in the table above. I used the formulae in Appendix B to calculate these estimates.

Appendix E.

Comparing the proportion of anglers by licence for each survey mode: First Contact and follow-up versus population







The mosaic plots above show the proportions of licence type in each survey compared to their respective population. A chi-square goodness-of-fit test was computed for each group (Gp1_electronic, Gp1_paper, and Gp2_paper) on the hypothesis that:

H_o: The proportion of 1-Day licensees, 8-Day licensees, and Annual licensees is the same in the first contact, follow-up contact, and population.

H_a: The proportion of 1-Day licensees, 8-Day licensees, and Annual licensees is not the same in the first contact, follow-up contact, and population.

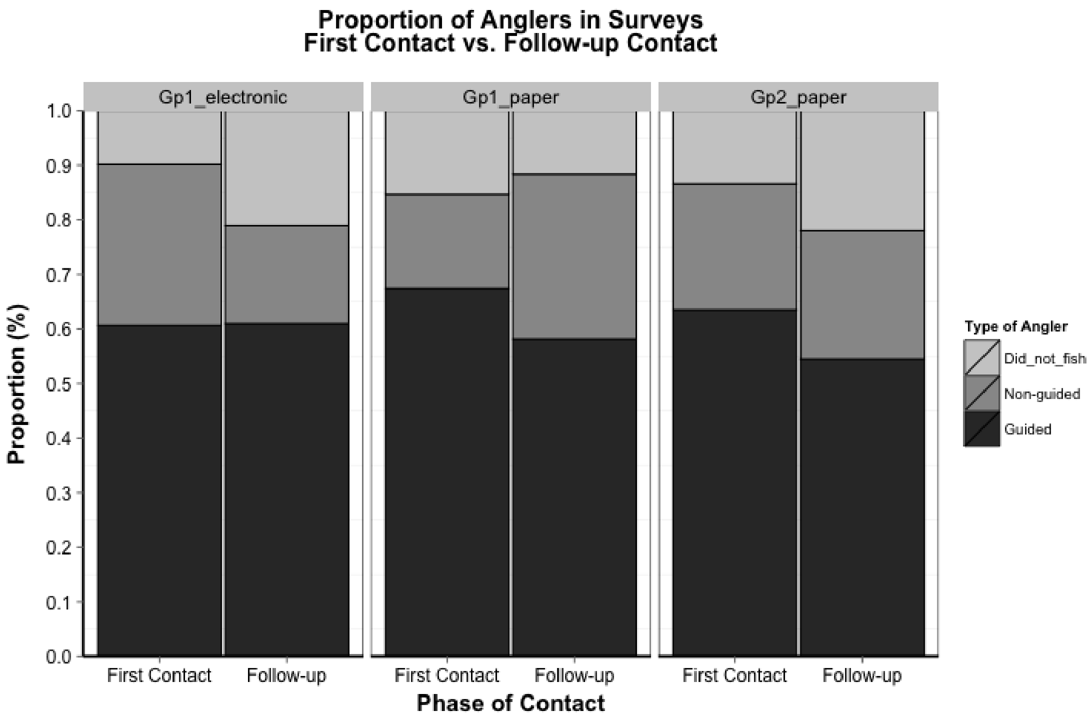
Results indicated that only Gp1_electronic (X^2 (4) = 26.4, p < .0001) exhibited a difference between the groups. The chi-square goodness-of-fit test did not provide evidence of a difference for Gp1_paper (X^2 (4) = 2.1, p = 0.7) or Gp2_paper (X^2 (4) = 3.5, p = 0.48). An inference can be made that Gp1_paper and Gp2_paper sample are representative of their populations. Licence type was deemed a good indicator variable because:

1. Licence type affects the fishing effort of anglers.
2. Licence type was also one of the only variable that had a known distribution in the sample and in the population.

Pertaining to Gp1_electronic, an inference can be made that anglers with an annual licence were more likely to respond to the first contact.

Appendix F.

Mosaic plot of proportion of anglers, by type of angler, in the first phase of contact versus follow-up contact



The mosaic plot above shows the proportion of anglers, by type of angler, in the first contact and follow-up contact for each survey. A chi-square goodness-of-fit test was computed for each group (Gp1_electronic, Gp1_paper, and Gp2_paper) on the hypothesis that:

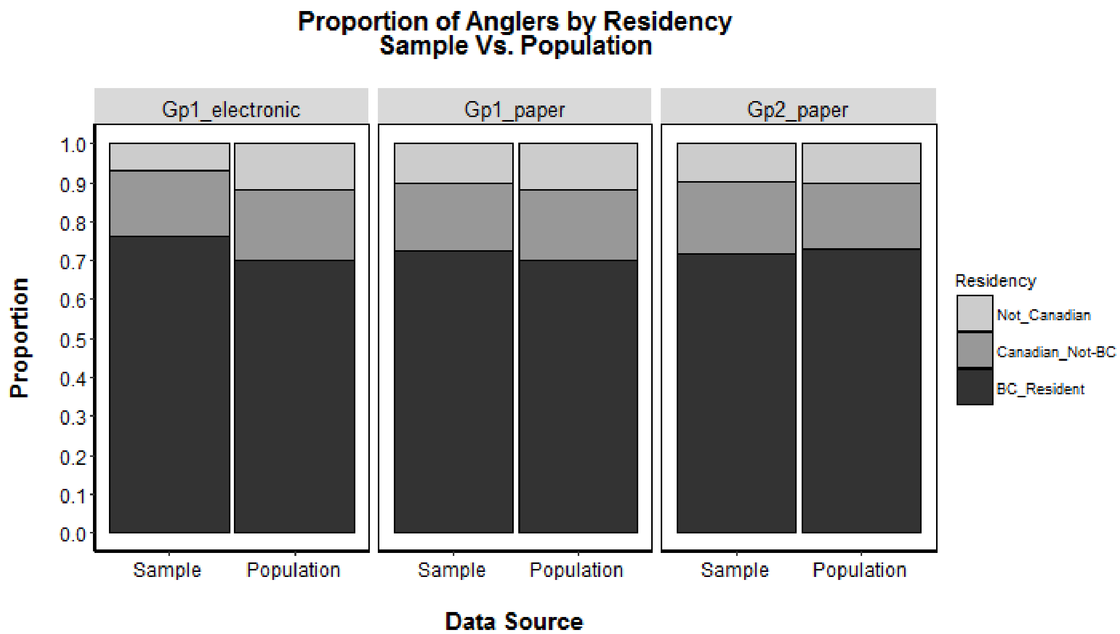
H_o: The proportion of guided anglers, non-guided anglers, and anglers who did not fish is the same in the first contact and follow-up contact.

H_a: The proportion of guided anglers, non-guided anglers, and anglers who did not fish is not the same in the first contact and follow-up contact.

Results provided evidence of a difference for Gp1_electronic (X^2 (2) = 29.6, p < 0.0001) and Gp2_sample (X^2 (2) = 7.2, p = 0.03). Anglers who did not fish responded in greater proportion to the follow-up survey. The test did not provide evidence of a difference for Gp1_paper (X^2 (2) = 4.4, p = 0.12). However, visual inspection of the mosaic plot indicated that proportions were likely different. The difference was evident for non-guided anglers.

Appendix G.

Mosaic plot of proportion of anglers, by residency, in surveys and their respective populations



The mosaic plots above shows the proportions of anglers, by residency, in the responding sample of each survey and their respective populations. Data are only for the first contact. Data from the telephone surveys are not included. A chi-square goodness-of-fit test was computed for each group (Gp1_electronic, Gp1_paper, and Gp2_paper) on the hypothesis that:

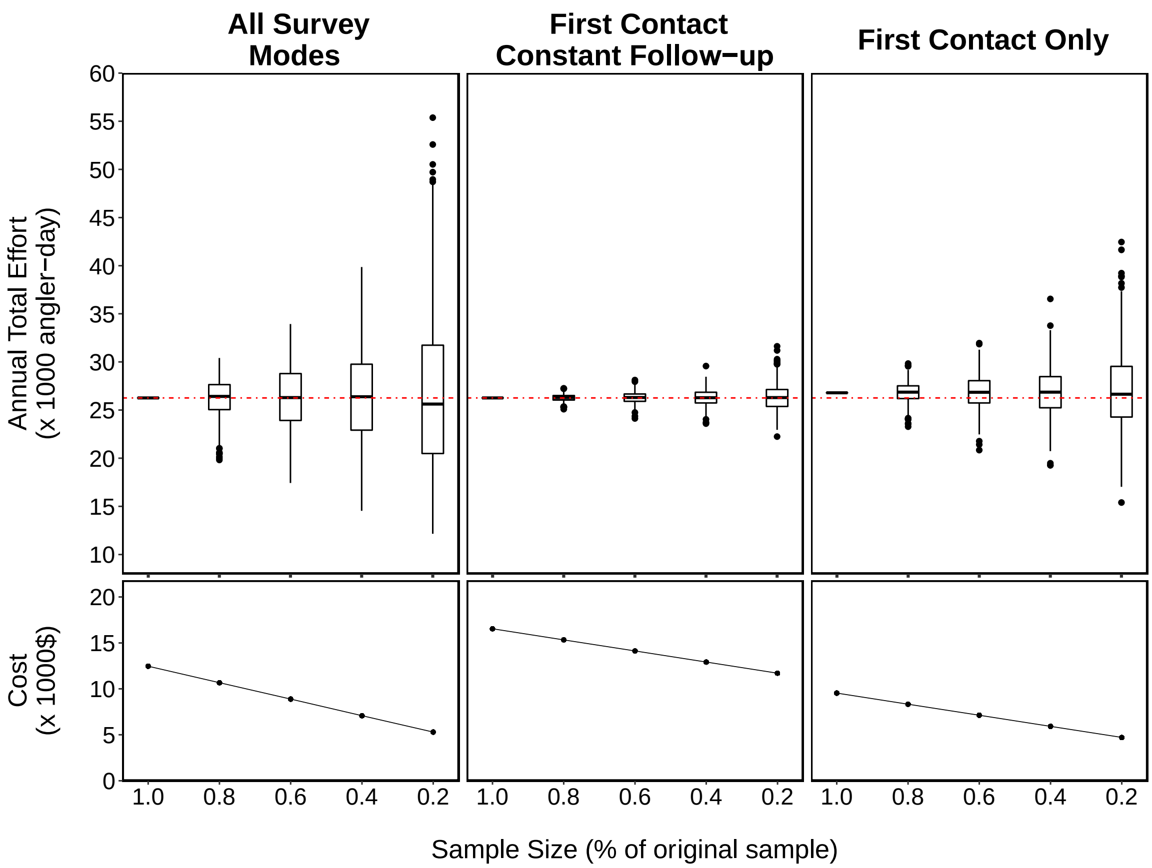
> H_o: The proportion of BC anglers, Canadian Not_BC anglers, and Not-Canadian anglers is the same in the first contact and the population.
>
> H_a: The proportion of BC anglers, Canadian Not_BC anglers, and Not-Canadian anglers is not the same in the first contact and the population.

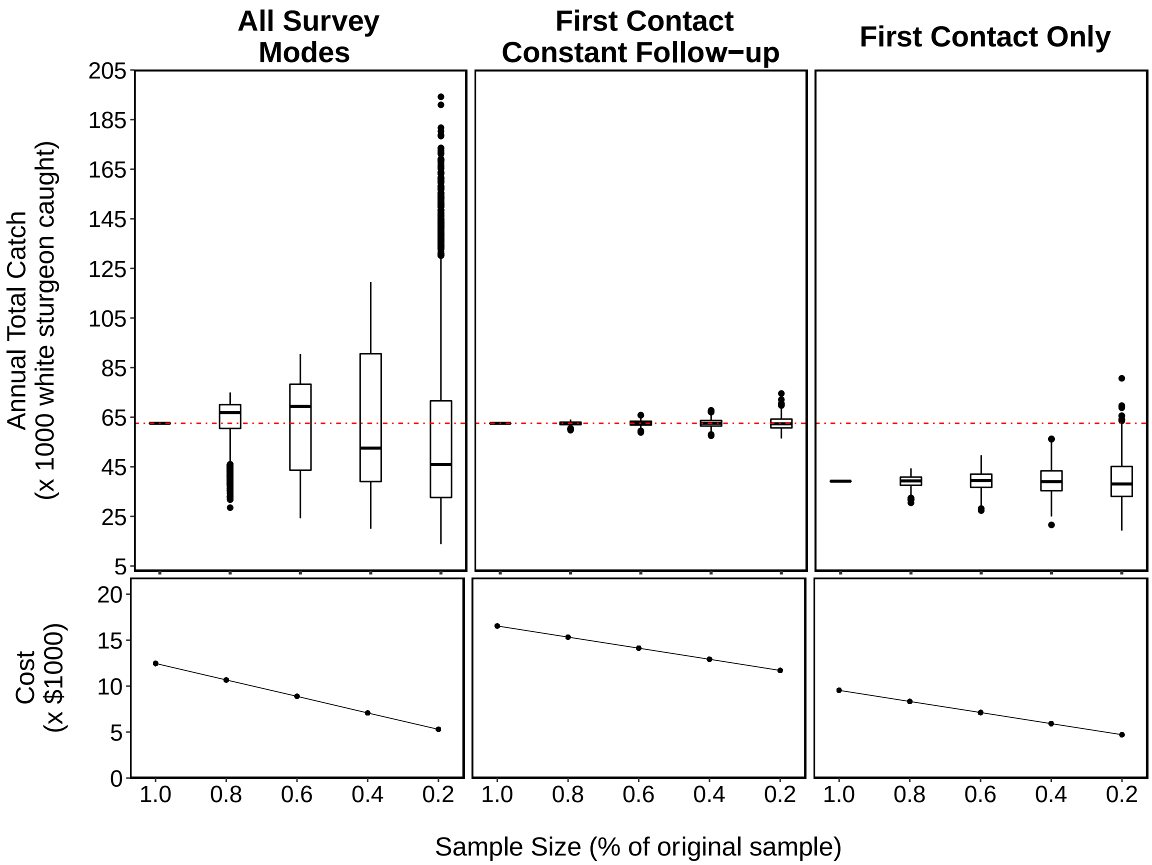
The chi-square goodness of test provided evidence that the proportions in Gp1_electronic were different for Group_1 population (X^2 (2) = 22.4, p < 0.0001). The chi-square goodness of test did not provide evidence for Gp1_paper (X^2 (2) = 1.6, p = 0.45) or Gp2_paper (X^2 (2) = 2, p = 0.37). Non-Canadian anglers were less likely to respond to the electronic survey.

Appendix H.

Variation of estimated population catch and effort as a function of sample size and survey effort

The graphs below shows the results of the resampling exercise for population total effort and total catch. Resampling was done without replacement. The formula expressed in equation 3 was used to calculate population totals. The cost of the surveys were broken down to cost per angler (total cost / sample size) for the regular mail survey and telephone survey. The cost of data entry was not included. The email survey had a fixed cost irrespective of sample size.





The variance of resamples in the left and middle panels increased drastically when the samples were reduced from 100% to 80%. This drastic increase in variance reflects the variance in the original samples. The mean of each sample was used to expand the sample totals to the population totals. Therefore, the population total was overestimated with small samples and a large amount of high catch/effort values since this will result in a large mean values. The panel on the right represents the resample with the sample size of the follow-up survey held constant. This yielded the most accurate estimates across all resamples. Even with 40% of the original sample (holding follow-up constant) the cost and RMSE are lower than with 80% of the full samples (i.e., left most panel). The resample exercise indicated that estimates are very sensitive to sample size when sample is reduced concurrently for all surveys. The sample size of the regular mail survey and email survey can be reduced by 60 percentage points but the follow-up survey must be included at the current sample size to yield accurate estimates. It must be noted that there is little value in reducing the sample size of the email survey because the cost is not dependent on sample size.