Development of User Perception Surveys to Protect Water Quality from Nutrient Pollution:

A Primer on Common Practices and Insights

Office of Water | EPA 823-R-21-001 | April 2021

EPA
United States
Environmental Protection
Agency

Disclaimer

This document provides information on user perception surveys. While this document cites statutes and regulations that contain requirements applicable to water quality standards, it does not impose legally binding requirements on EPA, states, tribes, other regulatory authorities, or the regulated community and its content might not apply to a particular situation based upon the circumstances. EPA, state, tribal, and other decision makers retain the discretion to adopt approaches on a case-by-case basis that differ from those provided in this document as appropriate and consistent with statutory and regulatory requirements. This document does not confer legal risks or impose legal obligations upon any member of the public. This document does not constitute a regulation, nor does it change or substitute for any CWA provision or EPA regulations. EPA could update this document as new information becomes available.

Contact Information

For more information, questions, or comments about this document, please contact:

Galen Kaufman
U.S. EPA, Office of Water, Office of Science and Technology
1200 Pennsylvania Avenue, Mail Code 4304T, Washington, DC 20460
kaufman.galen@epa.gov

Suggested Citation

USEPA. 2021. *Development of User Perception Surveys to Protect Water Quality from Nutrient Pollution: A Primer on Common Practices and Insights*. EPA 823-R-21-001. U.S. Environmental Protection Agency, Washington, DC.

Contributing Authors

Galen Kaufman. U.S. EPA, Office of Water, Office of Science and Technology, Washington, DC
Jacques L. Oliver. U.S. EPA, Office of Water, Office of Science and Technology, Washington, DC
Christopher Patrick. 2014-2015 American Association for the Advancement of Science, Science and Technology Policy Fellow
Natalie Spear. 2016-2017 National Oceanic and Atmospheric Administration Sea Grant John A. Knauss Marine Policy Fellow

Acknowledgments

The authors would like to thank the following people for help with background research, document development, formatting, and editing:

Jason Gershowitz and Erica Wales. Kearns & West
Jon Harcum, Susan Lanberg, and Dacia Mosso. Tetra Tech, Inc.

Additionally, the authors gratefully acknowledge the technical input and review from:

Association of Clean Water Administrators – Monitoring, Standards and Assessment Committee
Betsy Behl, Mike Elias, Vanessa Emerson, Claudia Gelfond, Sophie Greene, Deborah Nagle, Barbara Soares, and Dana Thomas. U.S. EPA, Office of Water, Office of Science and Technology
Steven Heiskary. Minnesota Pollution Control Agency
Scott Kishbaugh. New York State Department of Environmental Conservation
Tina Laidlaw. U.S. EPA Region 8
Jeff Ostermiller. Utah Department of Environmental Quality
Eric Smeltzer. Vermont Department of Environmental Conservation
James Summers. West Virginia Department of Environmental Protection
Mike Suplee. Montana Department of Environmental Quality
William W. Walker, Jr. Consultant

Contents

Executive Summary v

1.0 Introduction 1

1.1 Statutory Context 2
1.2 User Perception Survey Design 4
1.3 Methodology Used to Inform This Primer 5
1.3.1 Literature review 5
1.3.2 Interviews 5

2.0 Problem Formulation Development 6

2.1 Identifying the Problem 6
2.2 Identifying the Question to Be Answered 6
2.3 Is a Survey the Appropriate Tool? 8

3.0 Scoping, Designing, Conducting and Analyzing User Perception Surveys 8

3.1 Scoping: Assessing Resources, Opportunities, and Constraints 8
3.1.1 Existing resources and information 10
3.1.2 Types and amounts of water quality data 10
3.1.3 Expertise and staffing 11
3.1.4 Geographic scale 11
3.1.5 Reducing differences across waterbodies 13
3.2 Designing the Survey 14
3.2.1 Data quality objectives 15
3.2.2 Stakeholder engagement 17
3.2.3 Survey population/sampling frame 17
3.2.4 Sample size 19
3.2.5 Understanding sources of error 19
3.2.5.1 Sampling error 20
3.2.5.2 Sample selection procedures 20
3.2.5.3 Nonsampling error 22
3.3 Conducting the Survey 23
3.3.1 Survey options for interacting with the public 24
3.3.1.1 On-site survey 24
3.3.1.2 Online survey 25
3.3.1.3 Mail survey 26
3.3.1.4 Other survey modes 27
3.3.1.5 Mixed-mode surveys 27
3.3.1.6 Resource considerations for survey modes 27
3.3.1.7 Possible measures to address resource constraints 28
3.3.2 Survey questions 29
3.3.2.1 Picture selection 32
3.3.2.2 Beyond photographs 34
3.3.2.3 Auxiliary respondent data 34
3.3.3 Pre-implementation testing 35

3.3.4 Communication 36
3.4 Analyzing Survey Results 37
4.0 Ensuring Rigor in the Survey Process and Results 40
4.1 Quality Assurance/Quality Control 40
4.2 Maximizing Technical Rigor 41
5.0 Survey Design Scenarios 41
5.1 Examples of Considerations when Designing a Survey 42
5.1.1 Existing data 42
5.1.2 Existing program or stakeholder groups 42
5.1.3 Funding 43
5.1.4 Survey error 43
5.2 Scenario 1 43
5.2.1 Design 43
5.2.2 Method 43
5.2.3 Delivery 43
5.2.4 Analysis 43
5.2.5 Synopsis 44
5.3 Scenario 2 44
5.3.1 Design 44
5.3.2 Method 44
5.3.3 Delivery 44
5.3.4 Analysis 44
5.3.5 Synopsis 45
5.4 Scenario 3 45
5.4.1 Design 45
5.4.2 Method 45
5.4.3 Delivery 45
5.4.4 Analysis 45
5.4.5 Synopsis 46
5.5 Summary of Survey Design Considerations 46
6.0 Conclusion 47
Appendix A. References A-1
Appendix B. Interviews B-1
Appendix C. Survey Design Checklist and Questionnaire C-1

Executive Summary

Protecting surface water quality begins with observations of surface water characteristics. State and tribal water quality standards programs, acting under section 303(c) of the Clean Water Act,[1] often rely on discrete, quantitative measures of surface water's physical, chemical, and biological characteristics. Traditionally, these water quality observations have served as the analytical foundation for the development of numeric nutrient criteria. More recently, individual perception, or user perception, has emerged as an alternative measure of water quality. User perception of surface water quality, while related to traditional measures, is unique because of its integration of multiple environmental characteristics (e.g., color, water transparency, and biological features). Measures of user perception are also distinguished by their proximity to designated uses—which might explicitly refer to the protection of aesthetics—and narrative nutrient criteria—which might imply the prevention of adverse impacts to surface water aesthetics (e.g., no nuisance algal blooms). For these reasons, user perception surveys are an appealing tool for state and tribal water quality programs pursuing numeric nutrient criteria or translations of narrative nutrient criteria into numeric values for Clean Water Act purposes such as permitting and assessment. This primer provides an introduction to user perception survey design, implementation, and analysis for state and tribal water quality criteria regulators. It describes the methodologies associated with user perception survey design and the different ways to reach users when conducting surveys. It also addresses some of the important considerations in interpreting survey results when applying them to numeric nutrient criteria or narrative nutrient criteria translator development. Survey design scenarios and interviews with survey practitioners are also included to aid in visualizing user perception survey concepts.

[1] Title 33 of the *United States Code* section 1313.

1.0 Introduction

Aesthetic quality is an inherent and influential part of the human experience of water resources. Think about approaching your favorite body of water, whether it is a lake, reservoir, stream, river, estuary, or coastal marine water. The first things you perceive are how your surroundings look, smell, sound, and feel. These sensory experiences shape your mental and emotional responses. The intrinsic value of this psychological response is by itself important, as it influences your overall well-being and quality of life. This response also affects your willingness to spend time in and on that water, thereby affecting the potential enjoyment you might experience from recreational activities such as fishing, boating, and swimming.

Nutrient pollution is a widespread problem that affects the aesthetic quality of our nation's waters. It is an excess of nitrogen and/or phosphorus in a waterbody and is often measured as levels of total nitrogen (TN) and total phosphorus (TP). This pervasive problem in the United States has a strong influence on the characteristics that affect humans' perception of waterbodies and their experience in and on those waters. Waters clogged with growths of algae and plants, fish kills, and discolored and turbid water are among the many consequences of nutrient pollution that can affect how the public perceives and experiences a waterbody. These effects of nutrient pollution make waters unappealing to look at, clog fishing lures, cause long algae strands that can wrap around swimmers' legs and canoe paddles, decrease or eliminate fish populations that provide food and recreation, cover hard surfaces with slippery algal growths, and make swimming, recreational snorkeling, and diving difficult or dangerous by reducing visibility, among many other things.

The Clean Water Act (CWA) provides a mechanism states, territories, and authorized tribes[2] (hereafter, states) can use to protect aesthetic and recreational uses of waterbodies. Specifically, the CWA calls for states to create designated uses that take into consideration recreation in and on the water (33 U.S.C. 1313(c)(2)(A); Title 40 of the *Code of Federal Regulations* [CFR] 131.10(a)), and to develop criteria to protect those uses (33 U.S.C. 1313(c)(2)(A); 40 CFR 131.11(a)). Because the aesthetic condition of a waterbody directly affects how willing someone is to recreate in or on a waterbody, there is a clear connection between aesthetics and the recreational designated uses noted in the CWA. Some states have further recognized aesthetics as a waterbody quality to be protected by designating an aesthetics designated use. Because of these connections between aesthetic or recreational uses and CWA goals, the CWA can be used to develop criteria to protect these uses. User perception surveys are a tool that can be used to help quantify the concentrations of nitrogen and phosphorus that are protective of aesthetic or recreational designated uses, making surveys an effective method states can use within the CWA framework to develop protective water quality criteria to address nutrient pollution.

Quantifying user perception[3] through surveys is oriented around specific types of perceptions. Some survey tools attempt to quantify odor or lack of odor (as perceived by the user) at a waterbody. Surveys that focus on perception of odors are especially applicable when malodorous conditions are included as part of a state's narrative aesthetic standards or when a state receives complaints about the odor of a particular waterbody. Most state user perception surveys, however, focus on the visual aesthetics of a waterbody, which are estimated using either photographs or direct estimates of what a user visually perceives at the waterbody. This primer focuses on aesthetic user perception surveys as applied to

---

[2] The term "authorized tribe" or "tribe" means an Indian tribe authorized for treatment in a manner similar to a state under CWA Section 518 for purposes of Section 303(c) WQS.

[3] Use of the term *user perception* in this document, unless otherwise noted, refers to the judgment by a population regarding the usability of a waterbody based on a quality observable within the environment (e.g., green water or blue water, clear water or murky water). This differs from the concept of user perception in which the goal is to test the users' accuracy of perception and differences among individuals, which is related to a complex interface of human physiology and psyche.

respondents' perceptions of visual indicators of nutrient pollution in one or more waterbodies. A state surface water quality standards program (also referred to as a state), however, could consider using similar surveys to gather information on public opinions of other elements of aesthetic perception or on aesthetics related to visual indicators of other pollutants.

The purpose of this primer is to assist state water quality managers and scientists in developing user perception surveys to calculate target values for the derivation of numeric nutrient criteria. This paper describes some of the common practices employed in designing, implementing, analyzing, and interpreting user perception surveys. Because the information presented is drawn from expansive subject areas that are constantly evolving, this document is not designed to be an exhaustive treatment of all details related to surveys. It is meant to provide a general overview to help readers decide if user perception surveys are the right tool for them and, if so, to provide an introduction to key concepts and terminology. The document outlines many of the basic, but important, details water quality scientists and managers should consider, questions they should ask, and decisions they should make.

To create and conduct a survey most applicable to its situation and needs, it can be helpful for a state to perform additional research into survey methodology. To assist readers in further exploring topics presented here, this paper identifies resources in which additional information on each of the topics discussed can be obtained. Because individuals with survey expertise have practical knowledge in the field and familiarity with the latest standard practices, it is also recommended that they be consulted or brought in as members of the survey team.

This primer draws on current state water quality standards and regulations, the current literature related to user perception survey development and implementation, examples of state-implemented user perception surveys, and interviews with water quality scientists who have developed and implemented user perception surveys. This is intended to supplement existing U.S. Environmental Protection Agency (EPA) scientific approaches for numeric nutrient criteria development (USEPA 2000a, 2000b, 2001b).

1.1 Statutory Context

The CWA authorizes states to develop and adopt water quality standards, specifically designated uses and water quality criteria[4] to protect those uses (33 U.S.C. 1313(c)(2)(A); 40 CFR 131.11(a)). Recreational uses are specified in Section 101(a)(2) of the CWA, are among the most common state-designated uses (USEPA 2016), and include activities that occur in and on the water (e.g., swimming, boating, and fishing, respectively). Recreational uses are also directly affected by the aesthetic quality of a waterbody (NAS and NAE 1972) because the degree of recreational use–when, how long, and how often users interact with a waterbody–is strongly related to users' visual perceptions of aesthetic quality. Those perceptions help form users' evaluation of the condition of the water, and thus influence their willingness to recreate in or on the waterbody (Egan et al. 2009; Keeler et al. 2015; Smith and Davies-Colley 1992; Smith et al. 1995a, b; WHO 2003). Hence, whether or not a waterbody is meeting its recreational or aesthetic designated use can be determined by a better understanding of users' perceptions. With the use of surveys, user perception of aesthetic quality can be quantified and linked to specific biological and physical properties of a waterbody, which are often influenced by particular pollutants such as nitrogen and phosphorus (Figure 1) (Ditton and Goodale 1973; Nicolson and Mace 1975; Smith et al. 2015; West et al. 2016).

---

[4] Water quality standards are composed of three parts: designated uses, criteria to protect those uses, and antidegradation policies.

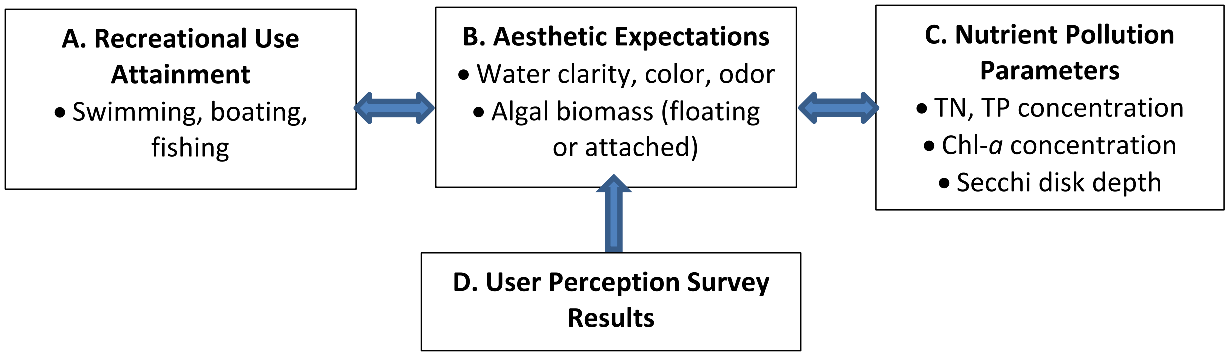

Figure 1. Conceptual model of the connection between recreational use attainment, aesthetic expectations, parameters related to nutrient pollution, and user perception surveys

States have been employing user perception surveys for many years to collect data on user perceptions and to inform the development of both narrative and numeric nutrient criteria (Brown and Daniel 1991; Heiskary and Walker 1988, 1995; House and Sangster 1991; Kishbaugh 1994; Smeltzer and Heiskary 1990; Smith et al. 1991; Smith and Davies-Colley 1992; Suplee et al. 2009). Some of the earliest examples from Minnesota and Vermont have served as models for surveys used by other states to correlate aesthetic perception of the suitability of water for recreation with water quality data (Heiskary and Walker 1988; Kishbaugh 1994; NYSDEC, n.d.; NYSFOLA and NYSDEC 2003; Smeltzer and Heiskary 1990).

Water quality standards in different states represent expectations of waterbody aesthetics in different ways. Some states have an explicit aesthetics designated use category for all or a subset of state surface waters (Environmental Protection Rule ch. 29A[5]). In some cases, states express aesthetic expectations implicitly as part of a recreational or general use designation (CEPA 2015; RIDEM 2009). In other cases, states might express expectations of waterbody aesthetics in narrative "free-from" criteria.[6] These criteria can also serve as the state's narrative criteria specific to nutrients or can relate to all pollutants (CDEEP 2013; MDEQ 2006; TCEQ 2014).

Although states express aesthetic expectations in a variety of ways, those expectations are often intended to protect recreational uses. Aesthetic qualities that relate to recreational uses include water color, water transparency or turbidity, the presence/absence of objectionable algal densities or nuisance aquatic vegetation, and the presence/absence of malodorous conditions. Aside from recreational uses, states also have aesthetic expectations in

User perception surveys have been used to examine water quality in:

- **Florida** (Hoyer et al. 2004)
- **Minnesota** (Heiskary and Walker 1988, 1995; Smeltzer and Heiskary 1990)
- **Montana** (Suplee et al. 2009)
- **New York** (NYSDEC, n.d.; NYSFOLA and NYSDEC 2003; Smith et al. 2015)
- **New Zealand** (Smith et al. 1991)
- **Texas** (TWCA 2005)
- **Utah** (UDEQ 2011)
- **Vermont** (Smeltzer and Heiskary 1990; VDEC 2016)
- **West Virginia** (RM 2012)

[5] Environmental Protection Rule Chapter 29A, Vermont Water Quality Standards. Vermont Department of Environmental Conservation.

[6] "Free-from" narrative criteria generally contain a statement describing the physical, chemical, or biological conditions that should not exist in the waterbody or waterbodies, that is, that they should be "free from."

their water quality standards to protect drinking water uses (e.g., taste, color, odor, or laundry staining) or fish consumption uses (e.g., fish taste, color, or odor) (ADEQ 2009; KDEP 2013).

Nutrient pollution contributes to many water quality characteristics most commonly associated with aesthetic and recreational use impairment (e.g., poor water clarity or discoloration; the presence of floating, blooming, and nuisance algae) (Figure 1, boxes B and C). The scientific literature continues to reaffirm this connection between nutrient pollution and adverse ecosystem impacts on recreation and aesthetics uses (Bricker et al. 1999; Dodds et al. 2008; Dodds and Smith 2016; Glibert et al. 2010; NRC 2000; Smith 2003; Smith et al. 1999; Vollenweider 1968). State water quality assessment programs also provide diagnostic evidence of these linkages, with state CWA Section 303(d) lists identifying nutrient-related parameters as the second most likely cause of recreational use impairment (USEPA 2016).

As shown in Figure 1, user perception surveys can be used to quantify aesthetic expectations and connect recreational and aesthetic use attainment to nutrient pollution. Results from user perception surveys (box D) can be either used to develop criteria for aesthetic expectations (box B; e.g., Suplee et al. 2009), or taken a step further and compared with known levels of indicators of nutrient pollution to determine concentrations of nitrogen and phosphorus (box C; e.g., VDEC 2016) that support the attainment of designated recreational and aesthetic uses (box A).

A theme throughout this paper is that there is flexibility when using surveys to allow each state to design a survey that reflects its statutory and regulatory circumstances.

1.2 User Perception Survey Design

Survey design and survey research are related fields through which best practices are continuously being refined as a result of advances in research, technology, analysis programs, and our growing understanding of human behavior. A state that elects to conduct a user perception survey of water quality should evaluate recent literature to ensure that it is using current best practices. There is a range of options states can use to capture user perception data and many factors that shape each survey process (see the examples in the inset to the right and section 3.0). Independently, each option can have different degrees of impact on the survey results. When combined, these options create the foundation for a valid, defensible survey process.

Examples of visual survey design options include:

- Targeting a specific waterbody type (e.g., lakes, wadeable streams) versus all waterbodies in an area (section 3.1.4)
- Surveying actual users present at a waterbody, past users, or likely/potential users (section 3.2)
- Conducting surveys using in-person interviews, mail, phone calls, or online questionnaires (section 3.3)
- Selecting the type and number of questions to get additional information about the survey population (section 3.3.2)

Despite this array of survey design options, effective water quality-based user perception surveys share common elements that facilitate the estimation of how users (or potential users) perceive surface water aesthetic quality. These common elements include problem formulation, consideration of opportunities and constraints, planning of survey design and implementation, analysis of survey responses, interpretation of the analysis, and methods for ensuring rigor in the conclusions.

This primer explores these elements and identifies common practices for conducting user perception surveys by drawing on past and present surveys as well as on the fields of survey design and survey research. Specifically, this document examines the following actions states might consider when developing a user perception survey:

- Deciding if a user perception survey is the appropriate tool

- Scoping the resources related to the survey
- Designing a survey
- Selecting methods to gather data from the public on aesthetic and recreation perceptions of water bodies
- Analyzing survey responses
- Implementing quality control (QC) and minimizing survey error

Section 5.0 provides several survey scenarios that can be used to support development of numeric nutrient criteria to protect aesthetics and recreational uses. States should consider which options are best suited to their needs in developing a user perception survey.

1.3 Methodology Used to Inform This Primer

Information for this paper was collected in two phases: (1) a literature review of published reports, papers, and journal articles related to existing surveys, and (2) interviews with experts and staff who have developed and implemented water quality user perception surveys. Appendix A provides a complete list of resources evaluated during the literature review, and Appendix B lists interviewees.

1.3.1 Literature review

The literature review informed many of the topics covered by this paper, including the existing practices related to conducting user perception surveys, best practices for survey design and implementation, and survey statistical research.

Sections 3.0 and 4.0 of this primer draw heavily on *Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Methods* (Dillman et al. 2014); *Survey Research and Analysis: Applications in Parks, Recreation and Human Dimensions* (Vaske 2008); and *Survey Management Handbook* (USEPA 2003).

1.3.2 Interviews

Eight interviews were conducted with individuals who have experience designing and implementing user perception surveys.

Interviewees included:

- Six individuals who each held a lead role in conducting surveys at the state level
- One individual from an EPA region who has worked with several states that have used user perception surveys to develop nutrient criteria
- One individual who worked as a consultant on several user perception surveys

2.0 Problem Formulation Development

Before designing and conducting a survey, the state should address some important questions. It is at this point, during problem formulation, that the state identifies what the problem is and what questions it wants to answer, then assesses whether a user perception survey is the most appropriate tool to use.

2.1 Identifying the Problem

In the context of survey design for development of numeric nutrient criteria or translator development, identifying the problem means identifying a realized or potential ecosystem problem caused by nutrient pollution that affects attainment of aesthetic and recreational designated uses in a waterbody. The evidence laid out in scientific literature provides a well-documented understanding of the link between nutrient pollution and designated use impacts.

In natural amounts, nutrients are essential to ecosystems, supporting the healthy growth of algae, aquatic plants, and microbes, which are the bases of aquatic food webs. Nutrient pollution of natural systems by excess nitrogen and phosphorus, however, destabilizes productivity and decomposition, causing a variety of problems.

The problems associated with nutrient pollution that affect aesthetic and recreational qualities of waterbodies are often caused by a shift in the composition of algal and plant assemblages toward species that produce nuisance growths, taste and odor issues, and toxins. Excess algal and plant growth are common symptoms of nutrient pollution. They reduce the attractiveness of water and can even physically constrain boating activities when algal and plant growths clog propellers or prohibit boat movement or access (Heisler et al. 2008; Smith 2003). Many nuisance algal species produce chemicals such as geosmin and 2-methylisoborneol, which affect taste and odor of water and fish tissue, and in turn affect recreation (Lopez et al. 2008). Some nuisance algal species also produce toxins such as microcystin (hepatotoxin), cylindrospermopsin (hepatotoxin/nephrotoxin), and anatoxin (neurotoxin), which affect human and wildlife health (Backer et al. 2003; Hilborn et al. 2014; WHO 1999). In some cases, the human health risks presented from exposure to algal toxins are high enough to warrant closing beaches.

These problems associated with nutrient pollution negatively affect the aesthetic qualities of waterbodies and can impact recreational uses by decreasing the public's willingness and ability to use waters for recreational purposes.

2.2 Identifying the Question to Be Answered

After identifying the problem to be addressed, the next step is to consider what must be known to address the problem. For numeric nutrient criteria or translator development, the ultimate question is, "What level of nutrients will support aesthetic and recreational designated uses?" To start answering that question, the first question to ask is, "What affects how aesthetic or recreational uses are achieved?" In the case of how nutrient pollution impacts aesthetic and recreational uses, a state wants to know how waterbody appearance affects public perception and at what level of nutrients.

Considering problem formulation in a manner similar to the one laid out in EPA's *Guidelines for Ecological Risk Assessment* (USEPA 1998) can help define the elements of a system a state is interested in. Using this method can help the state to clearly define the needed elements of the problem formulation by identifying management goals, assessment endpoints, and measures of effect. The paragraph below walks through an example of how to consider a problem formulation in a risk assessment fashion. To tailor this process, the state might select different management goals, assessment endpoints, and measures than those described in the example. It's worth noting that when developing criteria for recreational and aesthetic uses the process deviates from the conditions that predicate a true ecological risk assessment. In this case, the management goal is not to evaluate an ecological effect, but an effect on human perception and activity. This effect, however, is caused by a chain of ecologically based events, and it is, therefore, still helpful to consider a process similar to the one laid out in the EPA guidelines (USEPA 1998).

Management goals can often be found in state narrative criteria. For example, state narrative criteria often contain language such as "protection of recreational or aesthetic use from the adverse effects of nutrient pollution," "maintaining aesthetically pleasing conditions," or "maintaining recreational use." These sorts of statements describe the management goal(s) for the system. To select an assessment endpoint or endpoints, think about what ecological attributes are relevant to the management goal,[7] what in or about the system is valued,[8] and, in the case of nutrient criteria development, what attributes in the system are nutrient sensitive. In the case of the typical management goals above, the waterbodies of interest might be all lakes in a region and the attribute that is cared about could be the appearance of the lake. Because the state is interested in the aesthetic quality of the waterbody, it might measure the users' perception of the aesthetic condition of the water quality. For the purposes of numeric nutrient criteria development, additional measures are also needed. The first is a measure of response that affects the aesthetics of the waterbody, such as water column algal biomass (often represented as chlorophyll *a* [chl-*a*] concentration), percent benthic algal cover measures (typically represented as chl-*a* or percent cover), or algal concentration. Another component that will need to be measured is the stressor affecting the measure and assessment endpoint. In the case of nutrient criteria, TN and/or TP are the most important parameters, but the state could additionally consider the effect of confounding factors. A summary can be found below in Table 1.

Table 1. Example of using a risk assessment framework in problem formulation

Management Goal	Assessment Endpoint	Measure of Effect	Associated Measures Needed for Criteria Development	Stressor for Criteria Development
Protection of aesthetic and/or recreational use	Waterbody appearance	User-perceived water quality	Perceptible response variable - Algal density, concentration, coverage	TN and/or TP concentration

As noted above, chl-*a* and algal cover are commonly used measures for criteria-focused user perception surveys intended to relate management goals to stressors. Chl-*a* is a surrogate measure of algal biomass with a long history of application in water quality management. Representative values of chl-*a* that have been proposed or applied to evaluate and protect recreational uses for lakes include water column chl-*a* values of 2.6–7.0 µg/L in Vermont and 9–30 µg/L in Minnesota (varies by lake class) (Heiskary and Wilson 2005; VDEC 2016). For streams, an example of a chl-*a* recreational criterion is provided by Montana, which identified values greater than 200 mg/m^2 of benthic chl-*a* as undesirable and less than 150 mg/m^2 as desirable (Suplee et al. 2009). These values are consistent with previously published values of nuisance chl-*a* levels in streams (Welch et al. 1988). The other endpoint applied to streams has been percent algal cover, with recreational users in West Virginia considering greater than 20–25 percent algal cover unacceptable for recreation (RM 2012).

---

[7] What affects whether the management goal is attained? Often, in the case of nutrient pollution effects on aesthetic and recreational uses, that answer is algal or plant growth.

[8] Do people care if it's green? Do they care about the timing or location of the problem?

2.3 Is a Survey the Appropriate Tool?

A user perception survey can be useful to determine a nutrient concentration that is consistent with public perception of what is aesthetically desirable for its waters. That concentration can then be used to translate existing narrative water quality criteria or to support development of numeric nutrient criteria.

Once the problem(s) and question(s) to be answered have been identified, the state can then decide if a user perception survey is the most appropriate tool. The first step in determining whether a survey is the right tool for a state is to assess the current ecological landscape in the state and the state's waterbodies. If there is a strong visual change caused by nutrient pollution, a user perception survey can be a powerful tool to address questions pertaining to the change. The following are some helpful questions the state can ask prior to considering using a survey:

- Do changes in nutrient concentrations in a waterbody cause responses that can be visually observed?
- Are these visual changes consistent among the waterbodies in question (e.g., do all the streams of interest respond similarly)?
- Are recreational users able to detect gradations in these visual changes?
- Are these visual changes meaningful to recreational users or tribes?

If the answer to all of the above questions is "yes," then a user perception survey of the waterbodies can be a useful tool to assess how the population or specific users view aesthetic water quality. If the answer is no to one or more questions, or there is uncertainty of the answer, a user perception survey may still be the right tool, but additional work may need to be done to determine if it is.

3.0 Scoping, Designing, Conducting and Analyzing User Perception Surveys

The survey process as discussed in this section is divided into four basic phases:

- Scoping
- Designing the survey
- Conducting the survey
- Analyzing survey results

A state uses the survey design and implementation process to consider all of the pertinent aspects affecting the survey and determine the most appropriate methodologies to be employed.

3.1 Scoping: Assessing Resources, Opportunities, and Constraints

A user perception survey can take several forms, with each type having different information, effort, and resource requirements. Once a state has determined it wants to conduct a user perception survey, it can perform preliminary work to understand its existing resources and information, available expertise, scale of the effort, and waterbody specific considerations. This groundwork is important in helping the state determine if it has sufficient information and resources to address the type and scope of survey it is considering. The evaluations made in the scoping phase also inform later phases of the survey process and are often revisited as different options are considered.

A conceptual model for user perception survey scoping is provided in Figure 2.

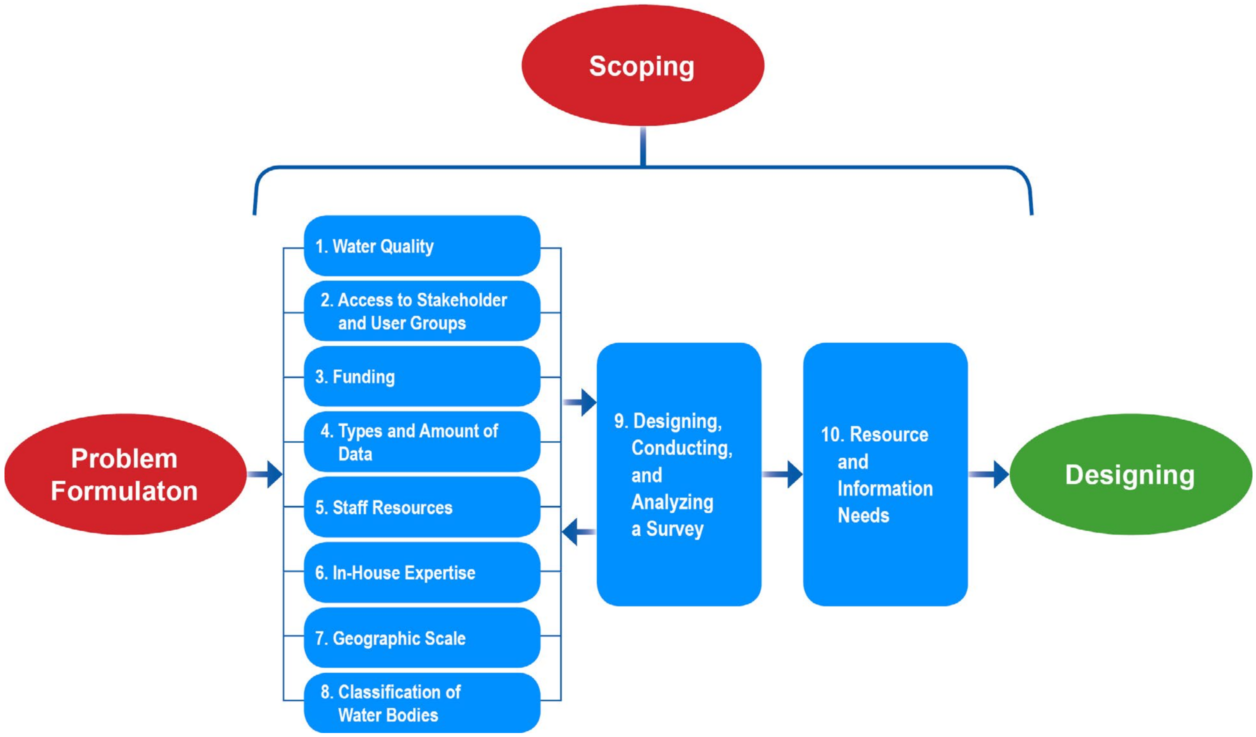

Figure 2. Conceptual model for user perception survey scoping

As shown in Figure 2, a state can ask scoping questions, as described in more detail below, to determine the breadth of a user perception survey.

1. What are the criteria and conditions of the waterbodies of interest?
2. What are your key stakeholder and user groups?
   What resources are available to help you contact them?
3. What level of financial resources do you have to conduct the survey?
4. What types of water quality data are available? What resources and information are available to support survey development?
5. Do you have enough staff time to conduct the survey?
6. Do your in-house staff have expertise in:
   - Survey design?
   - Survey research and statistics?
7. What geographic scale will be used to implement the survey?
8. Is it necessary or possible to classify the waterbodies?
9. How do the answers to the scoping questions affect the options available for designing, conducting, and analyzing a survey? As the survey progresses, decisions should be made in light of the scoping questions.
10. Based on the decisions made in the previous steps, what type and amount of resources are needed?

3.1.1 Existing resources and information

One important first step a state can take to inform survey design is to consider the level of resources available to develop and implement a survey. A state should take stock of existing resources and information at the beginning of the process to get an understanding of the systems of interest and the communities involved with those systems, and to take advantage of previous work. The state may want to consider the following aspects when assessing existing resources and information:

- Existing water quality criteria of the waterbodies in question
- Historic water quality conditions of the waterbodies in question
- Current water quality conditions of the waterbodies in question
- Available water quality or survey research conducted in the waterbodies in question
- Effects observed in the waterbodies in response to nutrient pollution
- Key stakeholder groups and their membership
- Methods used by other states to conduct user perception surveys
- Financial resources available to conduct the survey

Existing state criteria can provide a useful resource. A few interviewees found it helpful to include all or part of the existing regulatory language in the survey questions. Wording used in surveys developed in other states can also provide useful information. More about the use of existing regulatory and survey language can be found in section 3.3.2.

3.1.2 Types and amounts of water quality data

The types and amounts of water quality data that a state has available to compare to survey results can vary and shape the scope and form of a survey. States typically collect data on nutrient concentrations (e.g., TP and TN) and other factors that are not easily perceived, but nevertheless found in state criteria, such as dissolved oxygen (DO) levels. When employing user perception surveys, a measure of some easily perceptible quality of a waterbody that is responsive to nutrient pollution is also needed. Examples of those qualities include color, clarity of the water, type of plants or algae, and amount of plants or algae. These qualities can be quantified through a variety of field- and laboratory-based measurements, such as measures of water clarity (e.g., Secchi depth in lakes and estuaries, transparency tubes in streams/rivers), direct turbidity and color measurements, chl-*a* biomass and ash-free dry mass (water column in lakes; water column or benthic in streams, rivers, and estuaries), macrophyte biomass, visual plant or algal abundance estimates (e.g., percent cover, algal thickness, and algal filament length measures), and direct measures of algal/plant community composition (e.g., using microscopy for algae). Data for other parameters that might influence user perception of the quality of water (e.g., color or turbidity) as well as other parameters that might influence a waterbody's response to nutrients (e.g., temperature, pH, level of shading) are also useful during criteria derivation to clarify the relationship of nutrients to user perceptions.

User perception surveys can be carried out in lentic (lakes or reservoirs) and lotic (wadeable or nonwadeable streams) waters as well as in estuaries. The type of waterbody determines, to some extent, the water quality factors that are included in the user perception survey. For example, Secchi depth measurements are better suited for lentic waters, where biological expression of nutrient pollution is often in the water column as suspended algae, whereas in lotic waters nutrient pollution stress is often shown through benthic pathways, which may make benthic algae measures more appropriate (Smith et al. 1999). Since estuaries can be a mix of both types of environments, a variety of options is possible.

When selecting the water quality factors on which to focus in user perception surveys, it is important that the chosen factors are:

- Affected by nutrient concentrations consistently across all the waterbodies in question or the differences in how they are affected can be accounted for (e.g., through waterbody classification, see section 3.1.5),
- Affected by changes in nutrient concentrations in a visually perceivable manner, and
- Minimally affected by variables other than nutrient concentrations or the effects of other variables can be accounted for.

After determining the types of data available, it is helpful to inventory the data to determine their spatial and temporal characteristics. This inventory will be useful when considering if there are sufficient data to move forward with a survey, formulating data quality objectives (DQOs) (see section 3.2.1), determining where to conduct the survey, and selecting the relevant information to gather in the survey.

3.1.3 Expertise and staffing

Before designing a survey, a state should consider whether its in-house staff has the necessary skills and levels of expertise. If it does not, it should consider engaging outside experts.

> **"Getting the right people together with the right expertise is pretty critical."**

Using staff with survey and analysis experience, hiring outside experts, or consulting academics with survey design expertise is critical to successfully conducting user perception surveys. According to one interviewee, "contractor selection and making sure you have the right group is essential" to a successful survey. Designing and implementing a user perception survey and analyzing the results may require skills and expertise that the state does not have in-house. One interviewee emphasized that, "social science is a science. I think most [environmental agencies] are full of engineers and people trained in physical sciences, which use different language and techniques. Getting the right people together with the right expertise is pretty critical." A social scientist, whether in-house or brought in from outside, can work with physical scientists to ensure both sciences are captured in the survey and the intended results are realized.

The following areas of expertise might be needed for survey design, implementation, and analysis:

- Biology and ecology
- Survey design and methodology
- Statistical research and analysis
- Stakeholder/public engagement

One interviewee recommended that, when possible, all parties should participate in a kick-off meeting, especially if outside personnel are brought in as part of the survey team. Even if outside experts are not used, the survey team will likely be multidisciplinary. An initial meeting or meetings can help to provide background on the role of physical and social science, so that all individuals have an understanding of different languages used by both fields. This interviewee noted, "It helped a great deal trying to make sure we had both the [ecological] science and the social science right." Engaging all members of the survey team prior to the survey design process will help ensure that the survey meets the state's needs regarding stakeholder engagement and analysis.

3.1.4 Geographic scale

Depending on how broadly the nutrient criteria will apply and the level of regional differences, a statewide user perception survey might or might not be the most efficient tool. Most of the research used to develop this paper focused on surveys that were conducted across an entire state, but a state may want to conduct a user perception survey on a smaller scale, such as a particular watershed or ecoregion. This might be more useful than a statewide survey if there are significant differences in a state's waterbodies. In this case, a targeted survey might be more meaningful.

Case Study: Minnesota's User Perception Survey

Background/Environmental Question: The waters of the state of Minnesota are grouped into one or more classes based on their beneficial uses (e.g., aesthetic enjoyment and recreation) and the need for water quality protection in the public interest (Minn. R. ch. 7050).[1] The Minnesota Pollution Control Agency (MPCA) determined that excess phosphorus in Minnesota's lakes can stimulate algal growth that leads to frequent and severe nuisance blooms and reduced transparency, which, in turn, can limit recreational use of the lakes (Heiskary and Wilson 2005). As a result, MPCA set out to find answers to the following questions to support development of nutrient water quality criteria to protect the state's waters (Heiskary and Walker 1988):

- What is the relationship between the level of phosphorus in a lake and the frequency of nuisance algal blooms and reduced transparency in that lake?
- How do lake water quality measurements relate to subjective classifications or nuisance ratings based on physical appearance?

Why this Approach Was Taken: Prior to conducting the user perception survey for developing nutrient water quality criteria, MPCA had experience working with lake stakeholders (e.g., lake associations, county groups, volunteers) during a study of lake water quality. From this experience, MPCA realized the importance of using an approach in its criteria derivation process that associated chl-*a* concentrations or Secchi disk transparency with users' perceptions of water quality (Heiskary 2017, interview; Heiskary and Wilson 2005). MPCA began to develop a user perception survey and discovered that Vermont had already developed a lake observer study for waters in that state (Heiskary 2017, interview; Garrison and Smeltzer 1987, cited in Heiskary and Wilson 2005). After review of Vermont's survey, MPCA decided to integrate the Vermont survey questions into their lake monitoring program (Heiskary 2017, interview).

Recognizing distinct regions in their state, MPCA considered that definitions of "acceptable" or "objectionable" lake water quality could vary regionally (Heiskary and Walker 1988). For example, a lake user in a region dominated by oligotrophic lakes would probably have much higher expectations (e.g., higher transparency and lower algal levels) than would a lake user in a region dominated by hypereutrophic lakes. This observation supported MPCA's determination that an ecoregional approach to developing nutrient water quality criteria was needed (Heiskary and Wilson 2005).

How the User Perception Survey Was Conducted: User perception surveys were initially completed by MPCA staff members monitoring 40 lakes in early summer 1987. As part of these surveys, they collected water quality measurements (phosphorus, chl-*a*, and transparency) to compare to subjective classifications or nuisance ratings (Heiskary and Wilson 2005). To make the subjective ratings, staff members were asked to select one of five ratings of physical appearance (ranging from crystal clear to severe scums) and suitability for recreational and aesthetic enjoyment (ranging from no problems to no swimming) that most accurately reflected their impressions of conditions at the time of sampling (Heiskary and Walker 1988). As part of a larger monitoring effort conducted after initial survey implementation, Minnesota volunteer monitors also completed the surveys. The volunteers were usually members of the public who either lived on the lake or visited the lake routinely for recreation (e.g., anglers) (Heiskary 2017, interview). They recorded physical condition, their perceptions of water quality, and Secchi measurements of the lake several times a month (Heiskary 2017, interview). Using volunteer lake monitors to conduct the study meant that survey respondents did not represent a randomly chosen sample of public opinion (Smeltzer and Heiskary 1990). Before using volunteer lake monitors, the state weighed both the potential for bias as well as the benefits volunteer lake monitors may provide in criteria development through their awareness and knowledge of the signs and effects of eutrophication (Smeltzer and Heiskary 1990).

[1] Minn. R. (Minnesota Rules) ch. 7050, *Waters of the State*. https://www.revisor.mn.gov/rules/?id=7050&format=pdf.

(continued...)

Case Study: Minnesota's User Perception Survey (continued)

What Was Accomplished: MPCA found that the methodology used to determine the relationship between user expectations and lake water quality measurements was an important tool in developing nutrient water quality criteria (Heiskary 2017, interview). Water quality measurements were cross tabulated against user survey categories to provide a basis for calibrating nuisance criteria on a statewide and regional basis (Heiskary and Wilson 2005). MPCA found that there was reasonable agreement between user survey results and water quality measurements and that those relationships provided a rational basis for setting phosphorus criteria or management goals related to aesthetic qualities (Heiskary 2017, interview; Heiskary and Walker 1988). MPCA also found that differences between regions and survey response were statistically significant and especially pronounced at the lower end of the survey scale (i.e., responses of "crystal clear" and "beautiful, could not be any nicer") (Heiskary and Wilson 2005; Heiskary and Walker 1988). By conducting the survey, MPCA enhanced communication with stakeholders, and survey results provided stakeholder groups with a good understanding of the phosphorus criteria and how they correspond to perceptions, especially related to their lakes. MPCA continues to use the surveys and finds that their volunteers continue to be willing to collect data and make observations, which speaks to the affirmation of the approach (Heiskary 2017, interview).

States should also be aware that survey results could reflect regional cultural differences in perception of waterbodies and should take those differences into account in the survey design, or conduct surveys on smaller scales to minimize those differences (Smeltzer and Heiskary 1990). For example, New York State has very clear, deep lakes to the north where there is a lot of wilderness, and shallow, greener lakes to the south where there is more development and agriculture. Survey respondents from the north are more sensitive to decreases in water quality than respondents from the south, who are more accustomed to lakes with higher algal growth. As noted by one interviewee, "In the agricultural areas, when the water goes from very green to light green, they consider that great improvement. Whereas in [less developed areas], people see light green as a huge impairment. This regional [difference] is very important because it reflects actual management practices." In this situation, it would be useful to consider performing a survey on a regional scale or making a point to account for regional differences in survey design and analysis. In addition, states should also be aware that survey results could reflect cultural differences in perception of waterbodies by subgroups, such as tribes, and those potential differences should also inform the survey design.

3.1.5 Reducing differences across waterbodies

When using user perception surveys to develop nutrient criteria, it is important for the state to apply those criteria only to similar waterbodies. Not all waterbodies may exhibit similar responses to nutrient pollution, even if they are the same category of waterbody. For example, within the category of streams, first and second order streams may respond to changes in nutrients differently than fourth and fifth order streams. Categorizing waterbodies helps to reduce variability in the results during later analyses and produce more meaningful results. For example, the Plains streams in eastern Montana do not exhibit algal growth consistently in response to nutrient pollution. As a result, the Montana Department of Environmental Quality (DEQ) conducted a statewide public perception survey, but did not focus on Plains streams in either the survey or the numeric nutrient criteria that were elicited from the survey results. Instead, Montana DEQ developed a separate set of nutrient criteria unique to the Plains streams that focused on nutrient pollution effects on DO and fish (Suplee et al. 2009). Likewise, the Minnesota Pollution Control Agency (MPCA) compared Secchi readings to user evaluations to determine that regional differences in user perception of lake water quality occur in Minnesota. Accordingly, they chose to categorize their waters by ecoregion for their survey (Heiskary and Wilson 2005). One interviewee advised that, "If [a state] has a common waterbody type, it is easier to ask and get these questions [about perception] answered."

3.2 Designing the Survey

Survey design and research is a complex and fluid field. This section provides a general overview of the different aspects a state should consider when designing a survey. This is not meant to be a complete guide. To create a survey that is most applicable to its situation, a state may need to perform additional research into survey methodology and include members with survey expertise on the survey team to ensure that the results are as robust as possible. Active collaboration with experts in survey design, application, and statistics in all stages of the process can help ensure a high-quality product.

A state can adopt several of the options explained below to develop a user perception survey. General information about the level of effort, relative cost, and statistical rigor of the survey results is provided for these options and also summarized in Table 3 and Table 4 at the end of the primer. Note that this paper makes no direct differentiation between staff time and time of outside experts; each state should conduct an analysis to determine overall level of effort and specific expertise necessary to carry out a survey specific to their needs.

The following discussion addresses some technical design elements of the survey that are considered outside of how to administer the survey to the public and later analyze the results. The topics addressed here interact with the survey implementation and analysis phase, but are more focused on how the technical elements of the survey are constructed to ensure it is technically sound.

A conceptual model for designing a user perception survey is provided in Figure 3.

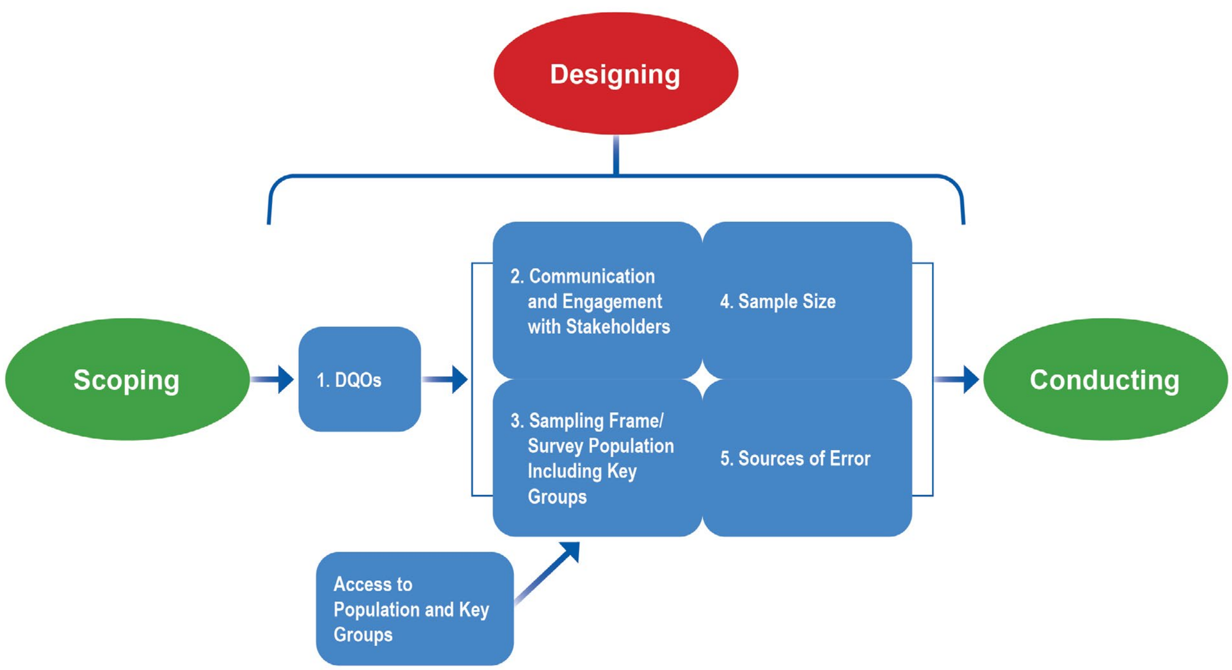

Figure 3. Conceptual model for designing the survey

As shown in Figure 3, a state can gather information, as described in more detail in the questions below, to determine how a user perception survey should be designed.

1. What are your DQOs?
2. Do you plan to engage stakeholders? If so, how?
3. What population do you want represented in your survey population (e.g., general population, key user/stakeholder groups, both)?
   What key groups do you want to target? This assumes that the state has the ability to access those groups.
   Do you have access to contact information for a general population survey?
   Do you have access to contact information for key user or stakeholder groups?
4. What is your ideal sample size?
5. What steps will you take to minimize survey error?
   What margin of error is acceptable for your survey?

3.2.1 Data quality objectives

During the survey design stage, the state can use the DQO process outlined in EPA's technical planning publication, *Guidance on Systematic Planning Using the Data Quality Objectives Process, EPA QA/G-4* (USEPA 2006), to help design a well-planned survey (Figure 4). For the collection of data, DQOs are needed to "clarify study objectives, define the appropriate type [and amount] of data, and specify tolerable levels of potential decision errors that will be used as the basis for establishing the quality and quantity of data needed to support decisions" (USEPA 2006).

When designing a user perception survey, the state can begin its DQO process by stating the problem(s), identifying the goal(s) of the survey, identifying information inputs, and defining the boundaries of the survey. Some aspects of this process are described in sections 3.0 and 4.0 of this primer. The state should also have some idea of the survey methodology and the types of analyses it wishes to conduct on the results. At this point the state can establish the DQOs, which are clearly stated requirements for the quality and quantity of new or existing data that will be collected. For the purposes of criteria development, these requirements are often expressed with some level of acceptable uncertainty. The last component of the DQO process is the development of a data collection plan that includes the "type, number, location, and physical quantity of samples and data, as well as the QA [quality assurance] and QC activities that will ensure that sampling design and measurement errors are managed sufficiently to meet the performance or acceptance criteria specified in the DQOs. The outputs of the DQO Process are used to develop a QA Project Plan and for performing Data Quality Assessment" (USEPA 2006).

Throughout survey design, implementation, and analysis, the state should document the survey process with explicit and detailed rationale for key decisions. This provides a vital record that can help defend the survey process against challenges.

Additional information on QA and QC considerations of DQOs is provided in section 4.0 of this primer.

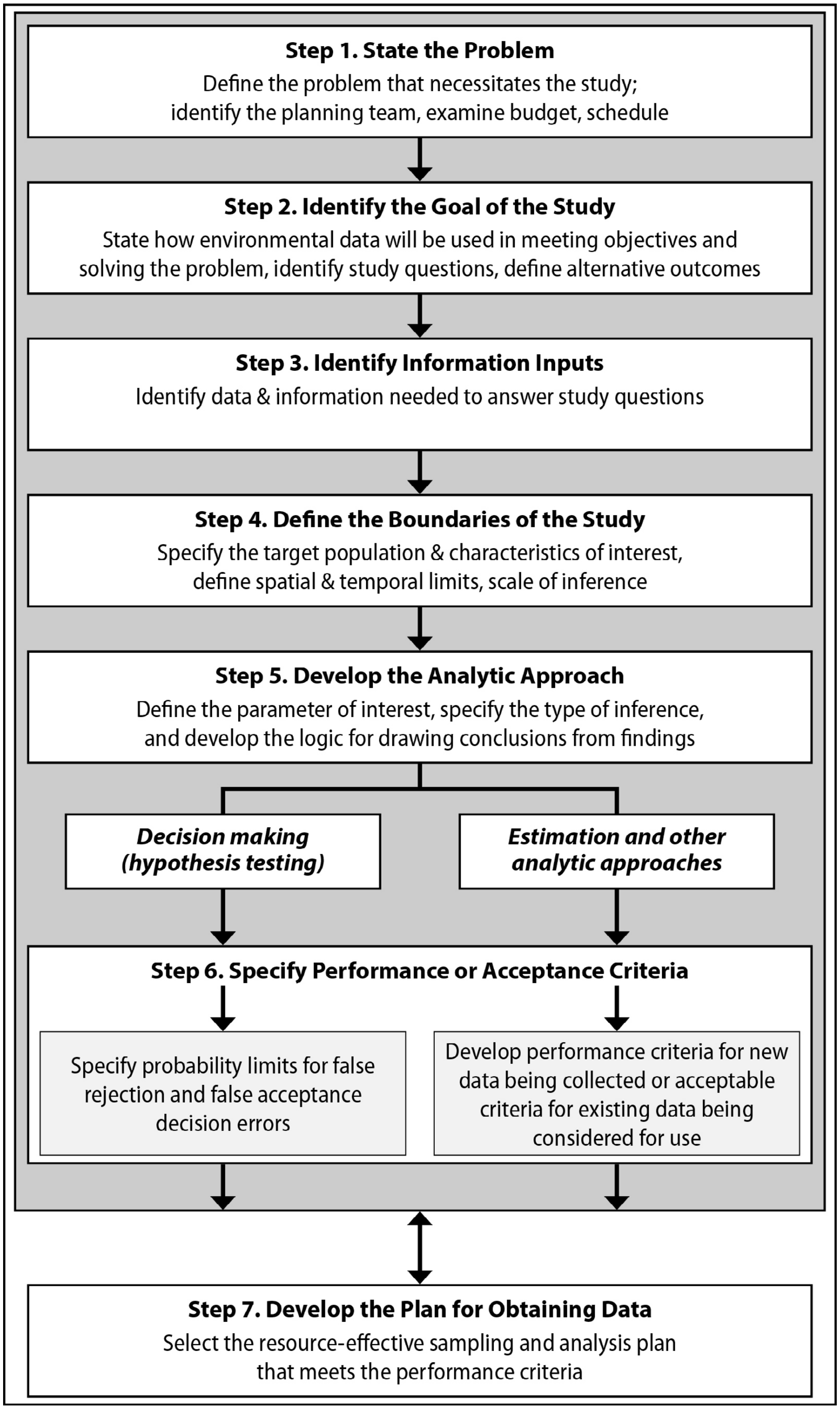

Figure 4. The process to develop DQOs (USEPA 2006)

3.2.2 Stakeholder engagement

> **Continued engagement throughout the survey establishes, grows, and reinforces trust between the stakeholders and the state.**

Stakeholder engagement throughout the survey process is a critical component not only for communicating information about the survey to the public and soliciting responses to the survey, but also for building support for the survey outcomes. States can develop support early in the process by engaging stakeholder groups in the design of the survey, which can reduce the risk of stakeholder challenges to the survey process. Stakeholder engagement can be particularly helpful if a state believes that stakeholder assumptions, such as a perception that citizens are very unhappy with water quality, will be challenged by survey results.

Stakeholder engagement also creates an opportunity for states to conduct focused analyses to understand and address specific perceived concerns. Engaging specific subgroups or identifying subgroups in the survey data provides an opportunity for more detailed analyses of these subgroups. For each subgroup of interest, the state should elicit responses from that subgroup to maintain the level of statistical significance and margin of error desired. In addition to getting buy-in to the process, stakeholder engagement also garners responses, which increases the meaningfulness of analyses.

Whenever a state decides to engage the public in some manner, it is helpful to consider the community it wishes to reach, including specific groups that:

- Have a financial interest in nutrient criteria (e.g., water treatment plants);
- Have an interest in water quality (e.g., recreational users);
- Make up a significant portion of the community;
- Are, or potentially could be, affected by environmental justice issues;
- Are vocal or politically influential; or
- Have historical or cultural connections to the waterbody or waterbodies of focus in the survey.

An extra effort to engage those groups could ensure that they respond and provide feedback, which will contribute to the success of the survey. If there are specific subgroups of the general population in which the survey design team is interested, then recording whether survey respondents fall into one or more of those groups and including them in the survey analysis will help ensure that they are represented. This might include documenting any demographic information, including underserved and minority groups that have access to the resource, as doing so will help address environmental justice considerations. Hiring translators might also be necessary to conduct surveys or focus groups.

During the design process, the state could even engage key stakeholder groups whose particular buy-in would be influential to survey success. While conducting additional outreach to key groups requires extra resources, the effort adds additional robustness to the final results of the survey and any management decisions informed by the survey.

3.2.3 Survey population/sampling frame

The survey population should be representative of the target population. As used in this document, the target population is often the general population that resides in a geographically defined unit (e.g., watershed) or political unit (e.g., county or state). In addition to ensuring the target population is represented, it is also important to take into account any subgroups of interest, such as the users of recreational waterbodies or specific user groups (e.g., anglers, beach users, tribal members). Information about the general community can be acquired from census records, and a randomized sample can approximate that population's make-up.

A sampling frame is a listing of the population that can be sampled and includes elements such as telephone numbers, addresses, and other characteristics that may inform the sample design, such as watershed, county, and distance to nearest waterbody. The elements of a sampling frame are called the sampling units. The choice of sampling frames and the steps taken to ensure their completeness and accuracy affects every aspect of the sample design (USEPA 2003). An ideal sampling frame should:

- Fully cover the target population;
- Contain no duplication (i.e., members of the target population are not represented more than once);
- Not contain elements that are not members of the target population (e.g., some voter registration rolls or telephone records might inadvertently include people who have moved out of the target area or are now deceased);
- Contain information for identifying and contacting the units selected for the sample; and
- Contain other information that will improve the efficiency of the sample design and the estimation procedures.

Using stratified sampling (in comparison to simple random sampling) when surveying over a broad geographic scale could improve sample design efficiency by producing estimates with smaller sampling errors (USEPA 2003). This could be useful in instances in which there might be reason to suspect that relatively poor or unique aesthetic waterbody conditions are more likely to occur in some parts of a state. If such regional differences are anticipated, geographic stratification could be used to select the survey sample to ensure that respondents who use waterbodies located in geographically representative parts of a state are included in the survey population. Sample selection procedures are discussed in more detail in section 3.2.5.2.

One way estimation procedures can be improved is through the collection of auxiliary respondent data, including age, gender, marital status, ethnicity, education, income, how frequently a respondent visits a particular waterbody, and purpose of visits (e.g., fishing, swimming). These auxiliary data could be used to develop a ratio estimate for application to an independent source of data, such as U.S. Census data (USEPA 2003) to reweight sample results to account for over- or under-representation. Auxiliary respondent data are discussed in more detail in section 3.3.2.3.

If the state does not have contact information for the targeted survey population or a sampling frame readily available, it may be necessary to expend a large amount of time and resources to obtain it, which would impact the cost of conducting the survey. There are several avenues the state can pursue to acquire contact information for the target population, including voter registration, driver's license registration, and tax records.

It is helpful for the state to consider the limitations of the list chosen to accurately represent the subgroups it wishes to target. For example, if a state has low voter registration or if there is a significant user group with lower registration, then voter rolls might not be an appropriate contact list to use for a general population survey.

States can also survey specific user groups by either using existing lists of known recreational users or surveying at recreational waterbodies. If there are specific user or stakeholder groups the state wishes to target, it could leverage resources such as:

- Sporting or fishing license registration lists
- Membership rosters or mailing lists for various interest groups, such as recreational, environmental, or community groups
- The same lists as described above for the general target population, but filtered by ZIP code

One interviewee shared that his state also weighted its survey population based on frequency of use by anglers. Because of the detailed nature of the existing data to which the state had access, it was able to survey known anglers using a random selection process that gave more weight to individuals who fished more frequently. The same state also chose the waterbodies to survey through a random selection process that weighted sites based on the number of visitors.

When designing a sampling frame and target population it is helpful to consider the characteristics of each and how they will affect the survey design. For example, if minors under the age of 18 are included in the sampling, additional Institutional Review Board (IRB) approvals and parental consent might be needed (see section 4.2 for more on the IRB). If there is a tribe in the target population the state should also determine if the tribe has any reserved rights relating to recreation in the waters of interest.

3.2.4 Sample size

Since it is not practical from a resource standpoint to interview the population of an entire state or region, a portion (or sample) of the population is interviewed. The sample size needed for a particular survey is dependent on the general target population size, whether any subgroups are specifically engaged in addition to the general target population, the amount of variation in the characteristics being measured in the population, the level of statistical significance desired, the margin of total survey error that is deemed acceptable, the expected response rate for the survey population, and the types of analyses intended to be done with the results. The actual number of responses needed to achieve the desired level of accuracy, consistent with the objectives of the survey, should be calculated with the help of a statistician or other experts familiar with survey statistics.

To perform the most basic levels of analysis on survey results of the general target population, 200–400 responses should be sufficient for typical levels of statistical significance and margin of total survey error (Vaske 2008). If the state wants to compare responses of the general target population against subgroup populations, or subgroups against subgroups, more responses will be needed to achieve statistical significance.

When designing a survey, it is possible to back-calculate (based on the expected response rate) the number of surveys the state will need to distribute in order to achieve a desired level of statistical significance and a margin of total survey error that the state deems acceptable (see section 3.3.1) (Vaske 2008). In general, a statistical significance of 95 percent is considered standard. Lower levels could be acceptable pending unique considerations for the survey (e.g., related activities to engage stakeholders).

3.2.5 Understanding sources of error

Error will always exist in survey design, but its influence can be minimized by careful survey design, understanding sources of error, reducing those sources as much as possible, and correcting for their influence in the results. The survey team can consider the total survey and determine how to best minimize total error within the given time and resource constraints. Since there is a range of options for conducting user perception surveys, it is best to select the one that is the most acceptable and suitable for a given scenario.

To ensure statistical rigor, it is critical that the survey design team minimize the following four types of error:

- Sampling error
- Coverage error
- Nonresponse error
- Measurement (response) error

The latter three types of error are collectively known as nonsampling error and are discussed in section 3.2.5.3.

3.2.5.1 Sampling error

An acceptable level of sampling error for key statistics is considered during survey design. The information derived from the data collected from the survey population (the sample) will differ from information derived from data collected from the entire population using exactly the same methodology. The difference between these two sets of values for every statistic is called the sampling error.

Selecting the margin of error and desired level of statistical significance is a risk management decision and a tradeoff between the cost to implement the survey and the needed level of precision to make decisions using the survey results. Smaller margins of error and higher levels of statistical significance are often desired when there is a high cost to implement (or not implement) decisions that result from the survey or if a state is more risk averse. Smaller margins of error and higher confidence levels require larger sample sizes, however, and thus result in higher costs.

Several resources can be used to determine the most appropriate level of statistical significance and the level of total survey error appropriate to meet the state's DQOs. Survey research and statistical experts can be helpful and can familiarize the survey team with industry standards and current practices. The survey team should also review reports or published articles on how other states have conducted similar aesthetic user perception surveys, including information on sampling error. Two interviewees for this report indicated that they conducted a random sampling approach using a plus or minus (+/-) 3.1- and +/- 5.0-percent margin of error, respectively, with a 95-percent confidence interval (CI), for the design of their survey sample sizes. That is, there was a 95-percent likelihood that the results from these two state surveys were within +/- 3.1 and +/- 5.0 percent of the target population. Larger margins of error (e.g., +/- 10 percent instead of +/- 5 percent) and smaller percent CIs (e.g., 90-percent CI instead of 95-percent CI) would lead to smaller required sample sizes.

To develop criteria based on user perception surveys, the state also needs samples of water quality variables to compare to the survey results. In addition to considering the sampling error associated with the survey, it is also valuable to consider the sampling error associated with measuring water quality in a waterbody over time and space. Because variability in water quality can influence users' perception of the water, the state should sufficiently characterize the water quality conditions associated with levels of perceived aesthetic quality. One way this can be achieved is by collecting enough samples over time and space to achieve desired levels of statistical accuracy. If heterogeneous patterns in water quality are suspected, it might also be helpful to classify or segment waters into areas of more homogenous conditions. For example, if differences exist in water quality between nearshore and offshore, or between different stream orders, these could be separated in a survey (see section 3.1.5 for more information).

3.2.5.2 Sample selection procedures

Various methods are available for achieving a random sampling of the targeted survey population to reduce sampling error while reducing the number of respondents needed. As the name suggests, simple random sampling is the most basic approach for sample selection, as all members in the targeted survey population have an equal probability of selection. This approach, however, might not provide sufficient information on key population subgroups. Other methods, such as stratified sampling, involve dividing the target survey population into non-overlapping subgroups and then selecting a simple random sample from each one. Multistage (cluster) sampling uses primary and secondary sampling units (Vaske 2008). For example, a state might randomly select the streams (primary sampling unit) on which to perform an on-river survey and then interview all individuals (secondary sampling units) recreating over a 10-hour period.[9] While simple random sampling is the most basic sampling approach, the other methods above reduce costs by allowing for improved sample design efficiency by producing estimates with smaller sampling errors (USEPA 2003).

The survey team generally develops a sampling plan that provides complete specifications for the procedures to be used for selecting sample units from the frame. As described in greater detail in EPA's *Survey Management Handbook* (USEPA 2003), the selection procedures in the sampling plan should specify any tasks necessary to reorganize or otherwise refine the sampling frame prior to selection, such as:

- Determining sample size and survey population;
- Eliminating units that are not in the target population;
- Identifying steps that will be taken to screen out ineligible sampling units, obtain better addresses, and so forth after the initial selection is made;
- Determining whether simple, stratified, or multistage (cluster) random sampling will be used;
- Transforming information about individual units into measures of size, which is needed for proportional, stratified-random sampling
- Determining whether the selection of sampling units at each stage will be with equal or variable probability. If variable probability is to be used, the basis for assigning selection probabilities to individual units must be included.

Sampling plans take into consideration the estimation procedures that might be used to convert sample data into estimates for the population. The approach used for the estimations plays a role in determining the size of the sample. In addition, some estimates require the capture of certain data (e.g., respondent ZIP code and age, how frequently a respondent visits a particular waterbody and purpose of visits [e.g., fishing, swimming]) when the sample is selected, during the survey's data collection phase, or during the survey's processing phase (USEPA 2003).

Estimation procedures include applying weights to give greater relative importance to some sampled elements than to others, adjusting for nonresponse, and using auxiliary information from the questionnaires, sampling frames, or other sources. Survey analysts can assign weights to adjust for sampled elements for which the probability of selection is in some way unequal, eligible units for which no data were collected (total nonresponse units), and sampling units not included in the sampling frame (noncoverage errors).

When selecting samples, it is important not only to consider the human sample population, but also other variables, such as the sample of waterbodies. For example, inclusion of a nonrepresentative sample of a state's population of lakes or streams in the user survey sites could introduce errors into the results, depending on the methods of data analysis used. For example, nonrandom selection of lakes included in Vermont's volunteer monitoring program made it necessary to synthesize representative distributions of lake water quality variables from a separate, probability-based lake sampling program before analyzing rates of false positive and false negative use impairment determinations associated with numeric nutrient criteria (Smeltzer et al. 2016). When possible, representative sampling of lake and stream sites based on the statewide distributions of nutrients and/or other water quality variables would avoid this type of problem.

[9] More information about the different types of sampling methods is available in Chapter 8, "Survey Implementation, Sampling, and Weighting Data," of Vaske (2008).

3.2.5.3 Nonsampling error

In addition to sampling error described in section 3.2.5.1, there are three types of nonsampling error associated with surveys (Dillman et al. 2014; USEPA 2003; Vaske 2008):

- Coverage error, which results from interviewing ineligible units or failing to interview eligible units
- Nonresponse error, which results when no data or incomplete data are obtained from eligible units
- Measurement (response) error, which results from incorrect reports by the interviewer or the respondent

Coverage errors are the result of a survey that does not correctly represent the population it was intended to sample. They can occur when incorrect listings of households are provided for a mail survey or in-person interviewers are given incorrect stream locations to perform on-river surveys. These errors can often be attributed to oversights in survey design; however, in some cases, the interviewers or respondents might be responsible for coverage errors. For example, interviewers might send a survey to the wrong household by addressing the envelope incorrectly. Fraudulent surveyors might even make up the answers to a questionnaire for a member of a hard-to-reach population instead of obtaining data from the designated respondent (USEPA 2003). These errors can be minimized by acquiring or developing an accurate sample frame and appropriate training, screening, and monitoring of interviewers.

Nonresponse error occurs when one or more of the selected respondents do not respond to the survey. What is considered an adequate response rate varies by research concentration and method. Some consider a 50-percent response rate adequate, while other researchers consider a response rate of more than 60 percent to be adequate (Vaske 2008). Typically, states determine the minimally acceptable rate of response (target response rate) necessary to achieve the research objectives and acceptable level of sampling error. Some researchers contend that a response rate of less than 70 percent for specific population segments signals a red flag for the research (Vaske 2008). However, it can often be the case that the response rate is lower than 70 percent depending upon the method used to elicit survey responses. Some typical response rates as well as ways to increase responses are discussed in section 3.3.1.

Nonresponse is primarily a problem if the nonrespondents are not a random sample of the total sample. For example, it is common for certain demographics not to respond to certain survey types. Failure to acquire data from all user groups might result in biased survey results. It is a best practice to actively investigate the potential for nonresponse survey bias prior to implementation (Dillman et al. 2014).

Measurement, or response*,* error includes bias and imprecision associated with sampling methodology. This error can be caused by the following:

- Method used to obtain responses (e.g., in-person, phone, mail, online)
- Presentation of survey questions
  - Wording
  - Order
  - Context
- Language barriers
- Respondent disposition
  - Confusion
  - Ignorance
  - Carelessness
  - Dishonesty
- Instructions and training received by interviewer

- Location of interview
- Deliberate errors

Preparing standardized survey procedures, performing pre-implementation testing (including testing electronic survey platforms, when applicable), and ensuring that the standardized survey procedures are consistently followed when collecting survey information can reduce nonsampling errors. More information on methods that can reduce nonsampling error for individual survey modes is provided in section 3.3.1.

More information about the types of survey error is available in EPA's *Survey Management Handbook* (USEPA 2003, ch. 4); in Dillman et al. (2014, ch. 1); and in Vaske (2008, ch. 8).

3.3 Conducting the Survey

States can conduct surveys of members of the public in different ways. Each method carries its own set of pros and cons that a state should consider. These various survey techniques are not mutually exclusive, and a state could choose to use one or more approaches. This section briefly lays out different modes states can use to interact with the public and collect survey response data, how to develop survey questions and select photographs, types of information that could be collected about survey respondents, ways to pre-test surveys, and suggestions for post-survey follow-up with respondents.

A conceptual model for conducting a user perception survey is provided in Figure 5.

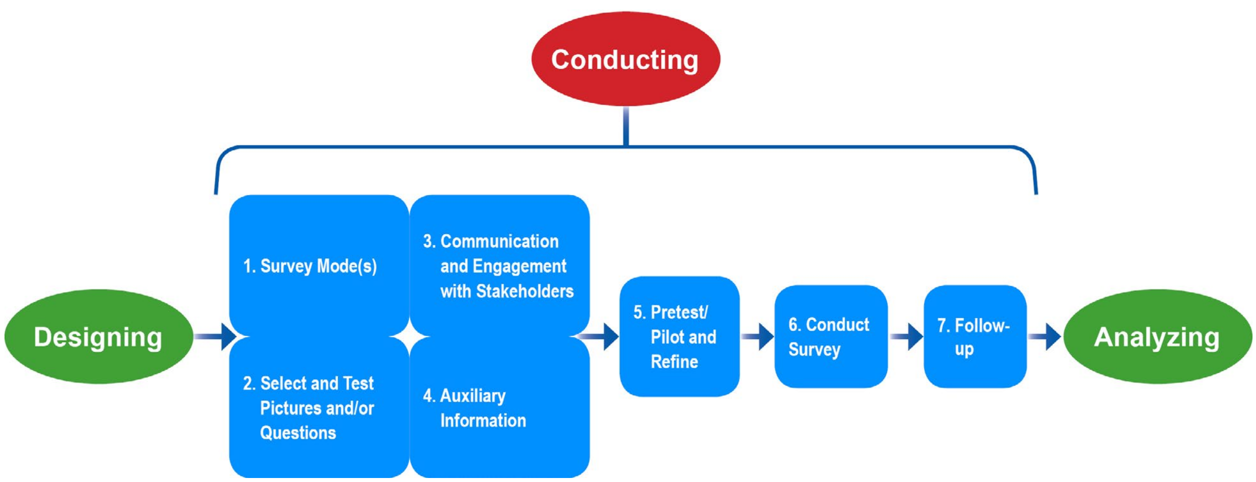

Figure 5. Conceptual model for conducting the survey

As shown in Figure 5 and described in more detail in the questions below, a state can gather information to determine how a user perception survey should be conducted.

1. What mode(s) will you use to conduct the survey?
   Looking at the comparison of survey modes, do you have adequate funding and staff resources to carry out the different modes you wish to use?
2. How will you select and refine questions and/or pictures used in the survey?
3. What are your plans for communicating with the public during and after the survey?
4. What types of information do you plan to collect about the demographics of the respondents?
5. How will you pretest or pilot your survey?
6. Do you plan to follow up with respondents after the survey and, if so, how?

3.3.1 Survey options for interacting with the public

> **Survey options for interacting with the public include:**
> - **On-site**
> - **Online**
> - **Mail**
> - **Phone**
> - **Mixed-mode**

3.3.1.1 On-site survey

On-site surveys specifically target users of the waterbody who are recreating in some manner on, in, or around it. Surveys conducted on-site at the waterbodies of interest could be done in numerous ways, including the following:

- Paper surveys distributed to respondents to complete on-site or at home
- Electronic versions of the survey given on a tablet/laptop computer or via a web link or mobile application to complete on-site or at home
- In-person interviews

Respondents could be asked to respond to preselected images or to conditions observed at the actual waterbody at the time of the survey. For surveys that ask respondents about current water conditions, responses are most meaningful when paired with water quality data taken at the same time, from the same site. This can be achieved with trained samplers (either staff or volunteer) taking measurements and water samples concurrent with on-site surveys.

> **Responses are most meaningful when paired with water quality data taken at the same time, from the same site.**

In-person surveys tend to have the highest response rate of the survey modes discussed here. The number of responses collected is dependent on the number of people at the site at the time the surveyor is out. As described earlier, the survey team needs to define the sampling frame carefully and determine what sampling units need to be surveyed. The state should have a good sense of the volume of visitors and peak days and times for individual sites to most effectively use staff resources. The survey team might be concerned with obtaining information from both the population that uses the waterbodies during the weekend and the population that uses the waterbodies during the week. Reweighting[10] can be used to handle some sampling bias after completion of the survey, but most of these considerations should be made when defining the sampling frame.

In addition to visitor records and logs that local and state agencies keep, the survey design team can use innovative technologies to determine which waterbodies are most popular. In one example, Keeler et al. (2015) analyzed geotagged photographs uploaded onto a photography-based social media site to obtain data on lake visits in Minnesota and Iowa.

If the on-site survey consists of one-on-one interviews, the interviewer(s) should undergo training on interview technique. The interviewer can be a potential source of survey error, and training can significantly reduce interviewer error (Vaske 2008). Studies have shown that the interviewers who received more training "produced lower item nonresponse, produced more complete recording of responses to open-ended questions, were more likely to read instructions and questions as worded, and

[10] The responses of the people sampled in a survey should be reflective of the overall population of interest. Sometimes this does not happen because of nonresponse or other biases that cause the responses of one or more groups to be over- or under-represented. Reweighting is an analysis method that is used post-survey to correct this problem. In reweighting, auxiliary respondent data can be used to assign a weight to each survey respondent's answers to make the overall response more similar to the population response if neither oversampling nor undersampling occurred. Reweighting is done by comparing collected auxiliary respondent data to known population data such as age, sex/gender, address, income, or other information. Weights are assigned to each respondent based on the difference between the auxiliary data for the sample and the population. More on auxiliary respondent data can be found in section 3.3.2.3.

were more likely to probe and to probe appropriately" (Dahlhamer et al. 2010). Helpful training topics include how to best approach a potential respondent, or not, if there are safety concerns, such as environmental conditions or safety issues with the potential respondent that would preclude conducting the interview. Training could also include how to explain the purpose of the survey and how to ask the questions in a neutral, nonleading manner (Vaske 2008). In addition, the interviewer should be professional, personable, and composed. To minimize response errors associated with in-person surveys, EPA also recommends that the interviewer establishes a good rapport with the respondent (USEPA 2003).

An on-site survey is not a random sampling of the population, which may or may not be an issue. Deliberate nonrandom sampling strategies can be employed if the state is particularly concerned with the opinions of specific user groups. Nonrandom sampling can, however, introduce coverage error in the survey if the state is concerned with a general population survey; those surveyed on-site are highly likely not to be representative of the larger population and are also likely not to include anyone who already thinks the water is too degraded to use for recreation. Whether a survey is random or nonrandom should be an intentional design choice. Associated assumptions should be examined to ensure consistency of the design with reality. Some states that have conducted user perception surveys wished to compare user groups with the general population and conducted on-site surveys in conjunction with a randomly mailed survey. One interviewee noted that the state "had wanted to do an on-river survey to get the opinions of people who were actually out recreating. But, because we knew the regulation would affect the general population, we also carried out a randomized mail survey of citizens."

While on-site surveys can be, but are not always, relatively inexpensive in terms of direct cost, they do require intensive staff time and energy. One interviewee's state, for example, hired an intern solely to conduct on-site surveys over the course of one summer at various sites across the state. The actual direct costs of conducting on-site surveys are determined by factors such as travel costs, training interviewers, and number of sites to be visited. States should consider all costs associated with conducting on-site surveys, including costs related to the sites surveyed, the frequency with which staff members are on-site conducting surveys, the method of survey delivery (e.g., interview vs. paper), and collation and tabulation of results.

3.3.1.2 Online survey

Today, it is easy to solicit and collect information online. There are numerous survey platforms that can facilitate states conducting user perception surveys online at very low cost and little staff time. The different survey platforms range in price, functionality, integration with other data tools, and degree of analytics available, and the state should choose a platform that matches its needs.

Response rates to electronic surveys, as well as mailed surveys (discussed below in section 3.3.1.3), are typically lower than response rates to in-person surveys, but can be improved through a variety of techniques. One strategy is to increase the respondent's interest in the topic or its perceived importance (USEPA 2003). The state can use different means, such as social and traditional media outlets, to increase the visibility and perceived importance of the survey with the public. The state can also partner with key user groups that can encourage their memberships to complete the survey. Prepaid incentives can also increase response rates. For example, in 2007, a survey of Washington and Idaho residents on quality of life that included a $5 incentive achieved a response rate of 70 percent (Dillman et al. 2014). Lastly, improved response rates can be achieved through survey follow-ups or additional contacts. The benefits of follow-ups can be twofold: (1) they improve response rate, and (2) they might provide some insight into nonresponse bias (i.e., hard-to-reach respondents might be a proxy for nonrespondents). Response rates improve substantially for each follow-up, although less for the last follow-up than for the first.

The biggest challenges with an online survey are controlling sampling and representativeness of the survey population, and verification of the sampled population. By not carefully controlling the survey population, the state introduces coverage and nonresponse error. It can be difficult to guarantee a truly randomized survey population, and very active and vocal stakeholder and user groups could be overrepresented. Additionally, individuals who learn about the survey via social media or a news outlet might be more likely to complete the online survey, or other survey mode, if they are already interested in the topic. Many online survey applications have means to limit respondents to the targeted population by requiring a specific passcode, multifactor authentication, invitation via a unique email link, or unique identifiers so there can be only one survey respondent per link or code. This could be used to help control the survey respondent population and minimize survey error. Following response collection, analysis can include a response audit for IP addresses or geolocation metadata to verify whether there is one survey respondent per link or code.

Online surveys can be a powerful tool that can engage a large number of people with very minimal effort on the part of the state. One interviewee who has overseen multiple user perception surveys noted that currently, states have not used digital surveys, but welcomed the idea of using them in the future. This person added, "I would not dismiss it. I would have bigger concerns if we didn't think about it." The state should carefully consider the trade-offs of using online surveys and how it can address error in both implementation and analysis.

3.3.1.3 Mail survey

A few states have mailed hard-copy survey forms to a randomized sample of the general population taken from a pre-existing, statewide voter registration list. These states asked respondents to fill out a paper survey and return it via mail.

While relatively low cost, this method typically results in a lower response rate than other modes. Interviewees stated they experienced response rates of approximately 20–30 percent and followed up with reminder mailings. Combining mailed surveys with an online survey and allowing respondents to return the survey via the mail or to complete an online survey could increase response rate.

Mailed surveys should include multiple communications, as this is the most effective way to increase response rate. Reaching out via multiple communications can also be useful for the other survey modes and should be considered as a means to increase survey response rate. Dillman et al. (2014) and Vaske (2008) suggest the following as possible multiple contacts:

- Prenotification letter
- Paper survey
- Thank you / reminder postcard or phone call
- Replacement paper survey

Without the follow-up reminders, response rates are typically 20–40 percent lower than when reminders are employed (Dillman et al. 2014).

An additional step that can be taken to reduce coverage errors for mail surveys is verifying street addresses for accuracy before surveys are sent to respondents.

The cost of a mailed survey increases greatly if the state does not have easy access to a contact list of the population it wishes to survey. It can spend large amounts of effort and money to get a general population contact list.

3.3.1.4 Other survey modes

This discussion does not include telephone surveys or other survey modes.[11] The user perception survey modes already discussed were the focus of our research. When designing a survey, the state should consider if other modes than the ones described in this primer might suit its needs more effectively. There also might be new technologies in the future that will allow for additional, not-yet-considered modes for conducting a user perception survey. One example of an innovative use of technology to examine user perception of water quality is discussed in a study conducted by Keeler et al. (2015). This study used geotagged photos from a social media site as proxies for lake visits by recreational users. They then compared the numbers of photographs taken at a lake with the lake's *in-situ* measured water quality, finding a positive correlation between the two. The use of mobile phone technology and applications is also currently being considered by researchers as a low-cost, convenient way to collect aesthetic and perception information in the future.

3.3.1.5 Mixed-mode surveys

It may be appropriate for a state to use a combination of the modes discussed earlier when conducting a survey in order to improve coverage of the targeted survey sample population, increase response rate, and/or reduce cost. This can be done either by offering multiple methods of responding to the survey (e.g., mailing out a paper survey as well as conducting on-site in-person interviews), but could also entail using different modes to reach out to the same population (e.g., mailing a physical letter containing a URL to an online survey, or following up to mailed surveys with a reminder phone call). Using a mixed-mode approach can increase survey response rate with very little cost; the survey design team should consider the most effective way to approach this.[12] Care is needed, however, to ensure that only one response is used per respondent and, where applicable, that reweighting is used to account for undersampling or oversampling of certain subgroups.

3.3.1.6 Resource considerations for survey modes

Conducting a user perception survey can sometimes be expensive and states may be constrained by time, staff resources, and financial resources. Table 2 provides relative guidelines on some considerations for three of the most commonly used modes of survey delivery—on-site, online, and mail surveys—that could impact whether a state elects to use each mode and to what extent.

---

[11] Research for this paper revealed only one state—West Virginia—that employed telephone surveys, although they combined them into a mixed-mode survey. More detail about West Virginia's survey is found in the case study in section 3.4. No additional surveys were found that included a major telephone component in conducting user perception surveys, so this primer does not go into great detail about telephone surveys. If a state wishes to explore conducting a telephone survey, it can consult Vaske (2008, ch. 8) and Dillman et al. (2014, ch. 8), which provide more information.

[12] More information about mixed-mode surveys is available in Dillman et al. (2014, ch. 11).

Table 2. Survey mode considerations

Consideration	On-Site	Online	Mail
Likelihood of nonresponse	Low	Low	High[a]
Cost per completed survey	High	Low	Low
Staff or outside expert time	High	Low[b]	Medium[b]
Anticipated response rates:			
General population	Medium	Medium[c]	High[d]
User or stakeholder group	High	Medium	Medium
Data collection time	Medium	Fast	Slow
Need for contact information	Low	Medium	High

Source: Vaske 2008, p. 126.
Notes:
[a] While relatively low cost, the likelihood of nonresponse is expected to be higher for mail surveys than for on-site and online surveys.
[b] Although it would require some time to develop an online survey and oversee its implementation, this type of survey will generally take the least amount of time to implement because many online platforms will gather all of the needed data and can include some level of data entry consistency. With mail surveys, there is more work involved with extracting the data and compiling data in one place.
[c] Online response rates would be expected to be medium because there are parts of the general population (e.g., elderly, populations with no internet access) that could be missed if online-only surveys are conducted.
[d] Anticipated response rates representative of the general population are expected to be higher for mail surveys than for on-site surveys and online surveys. On-site surveys might represent only a particular user or stakeholder group viewpoint. For online surveys, it can be difficult to guarantee a truly randomized survey population.

3.3.1.7 Possible measures to address resource constraints

It is important to note that none of the user perception surveys in our analysis were developed in isolation and that states often borrow survey language or design elements from one another to best suit their own purposes. When developing a survey, a state does not need to start the process from scratch. One interviewee recommended that a state can and should look to see what other states have done, particularly states of similar region, ecology, or political need.[13] This can save time and cost during the design phase. However, the state should consider the quality of the survey or the survey components it wishes to borrow. The state should copy a survey only if it is designed well and fits their needs.

A state can also identify the survey populations that are most important to consider when developing nutrient criteria and focus the survey on those populations. Reducing the size of the survey population has a direct impact on the resources needed to implement the survey. A state might be more interested in the opinions of recreational users than in those of the general population. Choosing to solely survey subgroups, however, increases the likelihood of coverage error and leaves the survey open to criticism of a biased respondent base. There are options for designing the survey to incorporate both general and subpopulations and decrease the survey effort needed. For example, the survey design team could elect to leverage one or more interested groups with which it is already in contact, such as volunteer water quality sampling teams, and have them complete perception surveys while in the field, while also surveying members of the public who are on-site (Hoyer et al. 2004; Smeltzer and Heiskary 1990). One interviewee noted that, by choosing that option, the state determined from analysis of the resulting data that “generally, those doing the sampling were more sensitive” to visual changes in water quality. The different levels of sensitivity of certain groups, such as volunteers, could bias survey result recording and interpretation of survey respondents. The unique experiences and skill sets of volunteers, however, could provide additional useful information for the survey. The tradeoff of reducing survey populations to achieve cost savings or other survey goals should be carefully monitored, and an analysis of results to determine response bias and validity is recommended.

[13] The inset in section 1.1 contains examples of places where user perception surveys have been used to examine water quality.

The survey design team might also want to explore technology solutions that can increase efficiencies in conducting surveys. For example, with the high volume of smartphone users today, it may be possible to decrease the number of on-site interviews needed by posting a Quick Response (QR) code linking to an online survey in prominent locations at popular waterbodies. States might want to explore options such as cost-saving measures, with the caveat that these tools should be pretested to ensure they achieve the expected number and types of responses.

3.3.2 Survey questions

User perception surveys of water quality have generally either used real-time conditions of waterbodies (asking questions about the visual aesthetics and recreational quality of the water as it is observed at the time of the survey) or used high-quality photographs of waterbodies.

The states studied tended to fall into two groups regarding the approach taken in their survey questions:

1. States that conducted their on-site surveys in conjunction with sampling and asked two questions. One question addressed the visual appearance of the water, and the other question explored the respondent's willingness to recreate in it.
2. States that presented survey respondents with a set of photographs of waterbodies with known chl-*a* concentrations and asked respondents to note whether each photograph depicted desirable or undesirable water quality.

States could also combine these two approaches, creating an opportunity to assess bias between respondents completing the survey on-site and those completing the survey in other settings. In any survey process, states should test assumptions for respondent bias whenever possible.

When developing a survey, it is crucial to first determine what study questions need to be answered to meet management objectives. For example, a state might find it important to determine whether the aesthetic conditions of a waterbody are considered to be of good quality (i.e., desirable) or would be considered barely acceptable.

After determining what overall questions need to be answered to meet management goals, states have several ways to approach the development of survey questions. Each approach affects the robustness of the survey process, especially when combined with other design decisions such as sample size or survey population. The process by which a state selects survey questions and pictures includes considering how to minimize survey error. A state could copy and customize another state's questions or images, or develop its own questions or images using techniques like A/B testing (Kohavi et al. 2009). Conducting a pretest or pilot among staff or in a focus group helps refine the survey and identify any issues that could affect responses once initial questions and photographs are selected (see section 3.3.3).

In developing the survey questions, one interviewee recommended using the existing narrative nutrient criterion to develop the survey, or a close paraphrase of the regulatory language. For example, if a state's water quality standards require that certain waters "exhibit good aesthetic value" or are "free from excessive algal growth," then it would be helpful to ask respondents to rate the aesthetic conditions by including word choices that include "good" or by directly asking whether algal growth is "excessive" at the time of observation. This way, an unambiguous yes/no determination can be obtained as to whether aesthetic impairment existed for the observer, making interpretation of the survey results much more straightforward when used to derive nutrient criteria or assess compliance with water quality standards. If the state is using existing regulatory language, it should consider that regulatory language might be difficult to understand in some cases, depending on wording used and survey audience. Pretesting or pilot-testing the survey can help to determine the appropriateness of regulatory language in the survey questions. If no existing narrative statement is being used and the survey is instead intended to inform the development of one, the language used in the survey could be used in the narrative criterion.

Case Study: **Vermont's User Perception Survey**

Background/Environmental Question: Vermont's water quality standards mandate that state waters be managed to achieve and maintain a level of quality that fully supports the specific designated uses of each water body. The types of designated uses that must be supported are aquatic biota, wildlife, and aquatic habitat; aesthetics; swimming and other primary contact recreation; boating, fishing, and other recreational uses; public water supplies; and irrigation of crops and other agricultural uses (VDEC 2016; Environmental Protection Rule ch. 29A).[1] As part of its effort to develop nutrient criteria to protect aesthetic uses, the Vermont Department of Environmental Conservation (DEC) set out to answer the question: "How can parameters such as TP, chl-*a*, or transparency be linked to impacts such as nuisance algae levels and recreational impairment?" (Smeltzer and Heiskary 1990).

Why this Approach Was Taken: Vermont DEC began developing numeric nutrient water quality standards in the 1980s. At that time, a state fish hatchery required a permit to discharge phosphorus into Lake Champlain, but Vermont DEC had no numeric nutrient criteria or total maximum daily loads to support permit limit development (Smeltzer 2017, interview). The department decided to measure aesthetics by implementing a lake user perception survey to link water quality measurements to recreational use and aesthetic conditions, for the purpose of protecting aesthetics (Smeltzer 2017, interview; Smeltzer and Heiskary 1990). After using the user perception survey to support development of numeric criteria for Lake Champlain and Lake Memphremagog (Smeltzer and Heiskary 1990), Vermont DEC then used the survey to collect information for Vermont "inland" lakes and reservoirs with surface areas of 20 acres or more to support nutrient criteria development for those water bodies (VDEC 2016).

How the User Perception Survey Was Implemented: Vermont DEC conducted the user perception surveys in coordination with the Vermont Lay Monitoring Program from 1987–1991 and again from 2006–2013 (VDEC 2016). The survey form used by the volunteer monitors consisted of part A, which asked monitors to rate the physical condition of the lake water, and part B, which asked monitors to rate the suitability of the lake water for recreation and aesthetic enjoyment (VDEC 2016). Ratings for part A ranged from "crystal clear water" to "severely high algae levels," and ratings for part B ranged from "beautiful, could not be any nicer" to "swimming and aesthetic enjoyment of the lake nearly impossible because of algae levels" (Smeltzer and Heiskary 1990). The volunteer monitors were asked to complete a user survey form each time they measured nutrient criteria variables (TP, chl-*a*, and Secchi depth) on their lake. This resulted in 5,073 individual survey responses for 87 different inland lakes (VDEC 2016). Vermont DEC found that this large sample was necessary to ensure coherent relationships between user responses and water quality variables (Smeltzer 2017, interview; VDEC 2016). Using volunteer lake monitors to conduct the study meant that survey respondents did not represent a randomly chosen sample of public opinion (Smeltzer and Heiskary 1990). Before using volunteer lake monitors, however, the state weighed both the potential for bias as well as the benefit volunteer lake monitors may provide in criteria development through their awareness and knowledge of the signs and effects of eutrophication (Smeltzer and Heiskary 1990).

What Was Accomplished: The lake user perception survey data have proven useful in a variety of lake management applications in Vermont, including numeric nutrient criteria derivation, statewide lake assessments, lake management goal setting, and wastewater discharge impact evaluation. For example, Vermont DEC used the relationships between user perceptions and TP, chl-*a*, and Secchi depth measurements to derive Vermont's nutrient criteria for inland lakes in a manner that minimized the risk of false positive and false negative use impairment determinations (Smeltzer et al. 2016). Vermont DEC found the survey results were useful in demonstrating to stakeholders how the phosphorus standards were derived (Smeltzer 2017, interview). Vermont DEC also found that using volunteer monitors to implement the lake user survey had two major benefits: (1) the volunteers collected field data, in addition to survey data, that was useful in developing nutrient criteria, and (2) working with volunteers allowed Vermont DEC to communicate information about lake water quality to lake residents because they were the ones participating in the survey (Smeltzer 2017, interview). Lastly, Minnesota has used Vermont's survey questions to develop its user perception surveys, indicating that the methods are useful for collecting information on user perceptions of aesthetics to support criteria development (Garrison and Smeltzer 1987, cited in Heiskary and Wilson 2005).

[1] Environmental Protection Rule Chapter 29A, Vermont Water Quality Standards. Vermont Department of Environmental Conservation. https://dec.vermont.gov/content/vermont-water-quality-standards.

In other cases, states could adopt the explicit language developed by other states for their surveys. Although the survey language being adopted might not be consistent with the regulatory use of the terms in the state adopting the language, using this approach allows for cross-state comparisons. When this method is being used, careful consideration should be given to the goals of the survey and the anticipated regulatory uses.

In an A/B test, users are randomly assigned to one of two groups (or variants): control (A) or treatment (B). Based on observations collected, an overall evaluation criterion is derived for each variant and analyzed to determine if the difference in the overall evaluation criterion between the A and B groups is statistically significant. For example, during an aesthetic user perception survey, the control (A) group could be shown pictures of local surface waters under existing conditions and asked whether the waterbodies are suitable for recreational purposes. The treatment group (B) could be shown pictures of waterbodies with lower algal growth than the control (A) group and asked the same question. If the A/B test is designed and executed properly, statistically significant differences in the overall evaluation criterion between the A and B groups could be used to show causality. Additional information on A/B testing is provided in Kohavi et al. (2009).

Example lake user survey questions used by Vermont and Minnesota

A. Please circle the one number that best describes the *physical condition* of the lake water today:
   1. Crystal clear water.
   2. Not quite crystal clear, a little algae visible.
   3. Definite algal greenness, yellowness, or brownness apparent.
   4. High algal levels with limited clarity and/or mild odor apparent.
   5. Severely high algae levels with one or more of the following: massive floating scums on lake or up on shore, strong foul odor, or fish kill.

B. Please circle the one number that best describes your *opinion* on how suitable the lake water is for recreation and aesthetic enjoyment today:
   1. Beautiful, could not be any nicer.
   2. Very minor aesthetic problems; excellent for swimming, boating, enjoyment.
   3. Swimming and aesthetic enjoyment slightly impaired because of algae levels.
   4. Desire to swim and level of enjoyment of the lake substantially reduced because of algae levels.
   5. Swimming and aesthetic enjoyment of the lake nearly impossible because of algae levels.

Source: Smeltzer and Heiskary 1990.

It is essential that all questions be “understandable to all potential respondents” and unbiased (Vaske 2008). Pretesting the survey is critical to ensure the questions are written free of scientific or policy jargon, in easily understood language, and are not leading or loaded. Questions should be simple, clear, balanced, and targeted to the information the state wants respondents to provide.[14] When designing questions, it is also important to include “not applicable,” “no opinion,” “other,” or a similar option with the responses the participants have to choose from to avoid a nonresponse issue when analyzing the results.

[14] General guidelines on writing good survey questions can be found in *Survey Research and Analysis: Applications in Parks, Recreation and Human Dimensions* (Vaske 2008, ch. 7).

The choice of open- or closed-ended questions is also a significant one. Open-ended questions allow for nuanced, qualitative insight into the user's perception of a waterbody. They can provide extra detail to supplement what might not be captured in closed-ended questions. However, because of the infinite choices of possible answers to open-ended questions, they can make for difficult quantitative analysis. If open-ended questions are used, it is crucial to accurately capture respondent input, so having audio recording can be helpful. It is important, however, to check to make sure what is legally allowed, and IRB input could be useful on this topic.

Closed-ended questions allow for easier statistical analysis of answers and can be more directly and simply used to derive quantitative criteria. Most surveys reviewed for this paper used closed-ended questions. Examples of closed-ended questions that can be useful for nuanced responses in water quality user perception surveys include a Likert scale and a rating scale, in which response to a question can be given on a graded scale (e.g., none, low, medium, or high algae cover).

3.3.2.1 Picture selection

We heard from several interviewees that photographs are a powerful tool for demonstrating the effects of different levels of nutrient pollution. For surveys that include images, the selection process for the photographs used is vitally important. Factors such as type of water quality factor represented, gradations of water quality shown in photos, photo quality, and standardization of photos can have large impacts on how water quality is perceived.

Photographs have to be carefully selected to be representative of a particular level of water quality (surveys that used images tended to focus on chl-*a* level) and at intervals at which there are discernible visual differences in the water quality variable for which the state has data. This means that with each photograph it is critical to have the associated field-derived measurements of the variables of interest, such as chl-*a* concentration or benthic algal biomass, to be able to determine gradations. One state that used an image-based survey had access to a large collection of waterbody photographs of known benthic chl-*a* concentrations. The survey development team selected eight pictures that were approximately 50 mg/m^2 chl-*a* apart and had similar perspectives and lighting (Suplee et al. 2009).

Another state used the same images for its image-based survey because of time constraints. While the photographs were sufficiently similar to the second state's own rivers, one representative of that state noted, "I wish we had more time. I would have insisted we had taken an extra year, if necessary, to get a photographer to make sure we had comparable photos, except for the variable [of interest]" to ensure the photographs were truly representative of the state's waterbodies.

A few interviewees identified various factors to consider when using image-based surveys. For example, it is crucial, although sometimes difficult, to standardize photographs in image-based surveys to minimize extraneous variables that are not directly related to nutrient pollution, but can impact user perception (Brown and Daniel 1991; Daniel and Vining 1983; Suplee et al. 2009). Following are some details to which to pay specific attention in each photograph and among the group of photographs used:

- Presence of background vegetation
- Presence of agriculture operations
- Proximity to urban settings
- Presence of litter
- Presence of fish or wildlife
- Quality and consistency of ambient light
- Color and hue of photographs
- Angle from which photographs are taken
- Flow level
- Water clarity

Case Study: Montana's User Perception Survey

Background/Environmental Question: Since 1955, the list of beneficial uses for Montana's streams has included public water supply, wildlife, fish and aquatic life, agriculture, and recreation (Suplee et al. 2008; ARM 17.30.601 *et seq*.).[1] In some parts of Montana (mainly in western regions), the most sensitive of these uses has been identified as recreation (Suplee et al. 2008). In wadeable rivers and streams, one way that the recreational uses and aesthetics can be impacted is by excess algal growth. It has been suggested that in excess of 100 to 150 mg chl-*a*/m^2 might constitute a nuisance (Horner et al. 1983, cited in Suplee et al. 2009; Welch et al. 1988, cited in Suplee et al. 2009); however, field validation was needed to determine what is considered a nuisance algae level to the public (Horner et al. 1983, cited in Suplee et al. 2009). The Montana Department of Environmental Quality (DEQ) set out to find an answer to the question: "What is too green for the public?" (Suplee 2017, interview).

Why this Approach Was Taken: In support of nutrient water quality standards development, Montana DEQ conducted a public opinion survey in 2006 to determine the level of benthic algae in wadeable rivers and streams that a majority of those surveyed considered undesirable for recreation (Suplee et al. 2009). The department used photographs of stream algae at different density levels in the survey to convey the environmental conditions that influence a person's perception of waterbody quality, or "key components," needed for participants to identify the levels of algae they would consider undesirable (Suplee et al. 2009). Montana DEQ selected two survey populations: the general public and on-river recreators (Suplee et al. 2009). The two populations were considered to be of equal importance because members of the general public are impacted directly by water quality standards, and recreators are most familiar with the water bodies since they use them. Montana DEQ also reasoned that recreators would provide the most accurate and consistent results (Suplee 2017, interview).

How the User Perception Survey Was Implemented: A by-mail survey of the general public was conducted by sending forms to 2,000 individuals randomly selected from Montana's Centralized Voter File (Suplee et al. 2009). An on-river survey of recreators was carried out in person at wadeable rivers and streams throughout Montana that had been selected based on known, statewide recreational use patterns (Suplee et al. 2009). Recreators at these wadeable rivers and streams included both residents and nonresidents. The mail and on-river surveys consisted of the same eight randomly ordered photographs of Montana rivers and streams, each depicting a different algae level ranging from less than 50 to 1,276 mg chl-*a*/m^2, staggered by approximately 50 mg chl-*a*/m^2 (Suplee et al. 2009). This concentration range was selected to cover the maximum range of benthic algae generally measured in Montana rivers and streams, and staggered with sufficient differences in algal density to permit visual distinction between photographs (Suplee et al. 2009). Montana DEQ tested the survey form and photograph sequence using generally accepted public opinion survey techniques (Suplee et al. 2009).

What Was Accomplished: In both the mail and on-river surveys, as benthic algal chl-*a* levels increased, desirability for recreation decreased (Suplee et al. 2009). Mean levels less than or equal to 150 mg chl-*a*/m^2 were found to be desirable by both survey populations (Suplee et al. 2009). Data analysis revealed no major differences in how the different survey populations perceived the amount of algae with regard to desirability for recreation. After implementing the survey, Montana DEQ had the information it needed to determine the amount of algae acceptable to the public, which allowed them to develop state water quality standards.

The survey results have been useful for communicating to stakeholders and the public how algae levels affect different beneficial uses and how the water quality standards were derived. Utah has also used Montana's methods to develop its standards, indicating that the methods are accepted by the scientific and management community and that the derived threshold is logical for recreational uses and aesthetics in western states. The recreational community was also generally pleased that Montana DEQ implemented these surveys in an effort to protect their waters, and the regulated community has not challenged the survey results (Suplee 2017, interview).

[1] ARM (Administrative Rules of Montana) 17.30.601 *et seq.* http://www.mtrules.org/gateway/Subchapterhome.asp?scn=17%2E30%2E6.

An additional factor to note is that image-based surveys do not capture all of the subtle environmental factors that cannot be transmitted visually (e.g., odor).

The state might also consider how it wants to account for possible visual variation in the aesthetic condition it is trying to portray. For example, there can be natural variation from lake to lake in how a chl-*a* level of 10 µg/L appears, depending on algae type, the distribution of algae within the water column (dots versus overall greenness), and other site-specific characteristics. To address this sort of variation, the state can decide to show the survey respondent the visual condition typical of 10 µg/L chl-*a* or, if it wishes to show different variations, it could present different conditions to assess differences in user response.

Aware of these concerns, one interviewee stated, “I agree there are inherent issues and bias in using photos. On the other hand, we’re asking questions about visual interpretation of something, so we have to find some mechanism to include photos people can respond to in order to get some gauge [of their perceptions].”

There are different ways to convey images in image-based surveys. Montana presented the photographs in a random order and asked survey respondents to consider each photograph independently. Respondents were asked “if the algae level was desirable or undesirable relative to his or her major form of river and stream recreation” (Suplee et al. 2009). West Virginia presented to respondents an online survey with a random photograph, and respondents were asked verbally over a scheduled telephone call if a picture depicted conditions that were acceptable or unacceptable. Respondents were directed to the next highest level of algae cover and asked about its acceptability, moving up until they found the level that was unacceptable (or until the highest level was reached). Conversely, if they found the first photograph to which they were directed unacceptable, they moved down in algae coverage levels until they found the level that was acceptable (or until the lowest level was reached). In this way, the survey recorded the highest level of algae cover that respondents felt was acceptable (RM 2012).

Utah provided images as a frame of reference for respondents, and asked individuals to compare the images to their own experiences (UDEQ 2011). In the context of using a user perception survey for criteria development, this question could be followed up with a question asking the user’s willingness to recreate in the waters pictured.

3.3.2.2 Beyond photographs

It is possible that future surveys could use means other than photographs to convey varying levels of water quality, such as video- or computer-generated images. If these other means are used, it is important that the survey design team carefully calibrate them, as they would for photographs, controlling other variables as much as possible.

There is some evidence that a nonstatic image is more sensitively perceived than a static image of the same scene (Hetherington et al. 1993). More thought and effort would be needed to control for different variables such as sound and other visual stimuli.

3.3.2.3 Auxiliary respondent data

The more data a state can collect about the survey respondents, the more types of analyses (e.g., cross-tabulations, analysis of variance, bivariate correlation and regression, and logistic regression) the team can conduct on the survey results. This supplemental information is separate from the survey questions related to water quality and includes personal data about individual respondents. One interviewee noted that collecting this extra data, which is seemingly unrelated to a respondent’s water recreation habits, can be helpful when performing various analyses.

Utah's User Perception Question

Algae are brownish green and short in length

Algae are dark green and long in length

Which of the following algae conditions did you usually see in late summer at this river?

a. Present algae are brownish green and short
b. Present algae are dark green and long
c. Algae are not present
d. Cannot see river bottom
e. Don't know

Source: UDEQ 2011.

Examples of auxiliary respondent data include age, sex/gender, address, marital status, race or ethnicity, education, income, religion, and other demographic information (USEPA 2003). Other useful information could include amount and type of participation in water-related activities, previous experience with water quality sampling or water quality issues, and location of interview. Some of these data are self-reported and included in the actual survey. Others are metadata and can be gleaned from ZIP codes (from mail surveys) or membership lists of different groups (from the mailing list of a known subgroup).

Collecting increasingly detailed information about respondents comes with some considerations that should be examined. Applicable laws and protections should be followed to protect respondent privacy, especially if any personally identifiable information is associated with respondent replies (it is generally easier and recommended that surveys are strictly anonymous). Additionally, there is a possibility that, as the detail of information being collected about respondents increases, they may become increasingly unwilling to participate or will skip certain questions, leading to nonresponse biases. Predesign research and pilot testing can help determine the appropriate types and amount of data to collect.

3.3.3 Pre-implementation testing

Although time constraints can be a concern during survey development, it is important to pretest surveys using focus groups to minimize survey error. Conducting a pretest or pilot is strongly recommended for new surveys as it helps refine the survey and identify any issues or problems that could affect responses. It is useful to pilot surveys both internally and externally as well as at the locations at which and in the manner in which they will ultimately be delivered. The state could leverage focus groups to pretest survey questions as well as images if an image-based survey is being used. A state could also conduct a full pilot of the survey on a small portion of the full survey population. The actual number of people used for the pilot will vary depending on time, resources, and feasibility, but the state should not be overly concerned with obtaining a large number. By watching individuals complete the survey during pre-implementation testing, researchers can also identify issues that could cause low response rate, skipped questions, and abandoned surveys (Vaske 2008).

Focus groups are useful tools through which to engage specific user or stakeholder groups. Focus groups convene a small number of individuals from a population of interest for a moderated discussion or direct questioning related to the survey. Generally, the survey team selects several relatively homogeneous groups of six to 12 participants from subgroups (e.g., anglers, swimmers, boaters) of the population of interest to participate in a focus group. The moderator focuses the participants on a few topics of special interest to the survey team (USEPA 2003).

Focus groups often help the survey team identify problems with the survey that might not be found otherwise. They also help identify the experiences of the selected segments of the population of interest, identify key concepts, help phrase questions so that they will be clear to all potential respondents, and evaluate drafts of survey questionnaires (USEPA 2003). One interviewee's state used focus groups to pretest the survey once it was initially designed and to retest it once it had been redesigned. This interviewee noted that the state "should have had some additional focus groups of random members of the public with directly affected stakeholders." Focus groups can also provide the state with valuable qualitative data to supplement quantitative data, but should not be used as a substitute for surveys as the results cannot be generalized to a larger population (Vaske 2008).

Keep in mind that focus groups cost time and money. For example, the state might wish to hire a professional facilitator to run the focus groups and direct the discussions if there is a need for neutrality or political complexity is involved with conducting a survey or developing nutrient criteria. In addition to the cost of running a meeting, the interviewee that used focus groups noted that participants received a small stipend to encourage participation. Focus groups take time to plan and run so the survey team should consider this when planning the overall survey development schedule.

3.3.4 Communication

As noted previously, stakeholder engagement and communicating the survey and survey results are integral elements to achieving buy-in and understanding of the process at the stakeholder and community level.

When developing a communication strategy around a survey, the state should determine the following:

- The target audience (see sections 3.2.2–3.2.3 for more information)
- At what stage in the process it will communicate
- How frequently it will communicate
- How it will communicate

The state could, for example, communicate to the following groups:

- The general public
- Survey respondents
- Specific user and stakeholder groups
- Policy-makers and legislators

Different audiences will require different levels and types of communication. To notify the general public, the state could use traditional and social media outlets to communicate information about an upcoming survey. This can increase the validity of the survey in the public's eyes and improve survey response rates (Vaske 2008). Additionally, the state can communicate with key user and stakeholder groups prior to releasing the survey, as described in section 3.2.2 of this paper, to increase buy-in to the survey process and request help to implement the survey if necessary (e.g., access to members' contact information).

While the survey is being conducted, the state can communicate with survey respondents multiple times. Doing so can greatly increase response rates (see section 3.3.1 for more information).

The state might also wish to follow up with survey respondents following their participation and after the completion of the overall survey process. If the survey is not conducted in-person, the state should, as good practice, follow up with a thank you/reminder note using the same method through which the survey was initially distributed (e.g., email or mail). This is an effective way to keep individuals engaged as well as to potentially solicit additional survey responses.

Follow-up communication can also be a means for the state to convey updates to the general public on the survey process, results of the survey, and how the survey is being used to inform the development of nutrient criteria. Follow-up could also be an avenue for public outreach and keeping in contact with an interested population.

After survey completion, the state will want to communicate the results to multiple audiences, including some that might not be receptive. One interviewee indicated receiving pushback to the idea of the validity of a social science survey from certain stakeholder groups with technical and scientific backgrounds (e.g., engineers). The state should consider its audience and anticipate any potential stakeholder challenges to the survey process and the survey results so that it is prepared to defend them. The state might also want to communicate the survey results to the respondents directly to keep them engaged and interested.

3.4 Analyzing Survey Results

If at all possible, the state should determine what types of analyses to conduct on the results prior to designing the actual survey. As a rule, any analysis conducted should be appropriate to the data. Different types of analyses require different types and amounts of data from the respondents and different sample sizes. This section lays out a general guide to the range of possible analyses that states can undertake.

Generally, higher levels of analysis require more data and variables, but have the benefit of adding increasing levels of precision to the survey results. If a state would like to perform higher level analyses, it should factor that in during the design phase when determining what additional data, such as background sociodemographic variables (e.g., socioeconomic information, race/ethnicity, education level) or content-specific variables (e.g., attitudes toward pollution, attitudes toward recreation), to collect from survey respondents. Also, during the design phase, it is helpful to consider using closed-ended questions to help analyze the data. For example, provide options when asking about race, or provide a gradient of choices using a Likert scale or similar rating scale when assessing attitudes.

As described in section 3.2 of this primer, the survey team generally determines the desired level of statistical significance and margin of error and develops a sampling plan before conducting the study. This sampling plan should include the analytical plan, part of which takes in to account anticipated response rate and the make-up of those responses, as well as estimation procedures that could be used if the response rate or response composition is not met after the survey is completed. The primary method for checking for nonresponse errors is conducting a nonresponse bias check. These checks are primarily implemented by following up with a subset of the sample population that did not respond to the survey. If the results from the nonrespondents are similar to the population of those who did respond, then the survey analyst could use reweighting factors or duplicate values reported by the sampled units to compensate for the nonresponses (USEPA 2003; Vaske 2008).

Case Study: West Virginia's User Perception Survey

Background/Environmental Question: West Virginia's water quality standards mandate that algal blooms or bacterial concentrations that might impair or interfere with the designated uses of the affected waters are not allowed in any waters of the state (47 CSR 2).[1] At a minimum, all West Virginia waters are designated for propagation and maintenance of fish and other aquatic life and for water contact recreation. When concentrations of nutrients are high, normal background levels of algae can grow into excessive blooms that interfere with the designated uses of a water body, including water contact recreation and public water supply (WV DEP 2017). As part of its effort to determine whether designated uses of West Virginia's surface waters are being impacted by algae levels, the West Virginia Department of Environmental Protection (DEP) set out to find an answer to the question: "What is the public's tolerance to various amounts of algae in West Virginia's streams and rivers?" (Summers 2017, interview; WV DEP 2017).

Why this Approach Was Taken: West Virginia DEP commissioned a user perception survey to determine West Virginia residents' tolerance of algae levels in the state's streams and waterways, and to evaluate the impact of algae levels on public recreational use. Seven photographs were selected to be used in the survey, showing various levels of algae cover, ranging from 4 to 65 percent, to assess each respondent's acceptance threshold for algae (RM 2012). West Virginia DEP reasoned that the general public was directly impacted by algae in West Virginia's surface waters; therefore, a random sample of state residents was selected from the West Virginia voter registration list as the population for the user perception survey. The survey was pretested before being implemented to ensure proper wording, flow, and logic (RM 2012; Summers 2017, interview).

How the User Perception Survey Was Implemented: In February and March 2012, West Virginia DEP engaged a contractor to conduct the user perception survey of West Virginia residents of at least 18 years of age. Respondents were initially contacted by telephone. If they agreed to participate in the survey, they were instructed to access a website that showed the preselected photographs of algae cover. If participants were not able to access the internet, they were mailed a printed packet. When participants were able to access the photographs, whether online or in printed form, they were interviewed over the phone to complete the survey. Respondents were assigned a random image as a starting point and asked to indicate if the level of algae cover shown was acceptable or unacceptable to them. If the first randomly selected image was deemed unacceptable, the respondent was shown the view of the next lowest level of algae cover, continuing to be shown lower algae cover levels until either an acceptable level or the lowest cover level was reached. If the first randomly selected image was deemed acceptable, the respondent was shown the next highest level of algae cover, continuing to be shown higher algae cover levels until either an unacceptable level was found or the highest cover level was reached. Respondents were also asked about their participation in aquatic recreational activities and their overall opinion of the amount of algae in West Virginia waters. Interviewers who performed the surveys were trained according to standards developed by the Council of American Survey Research Organizations. Also, a central polling site was set up to ensure that interviews were performed correctly (RM 2012).

What Was Accomplished: The results of the survey indicated that the majority of West Virginia residents would consider waters with more than 25 percent algal coverage to be unacceptable. Cross-tabulations of tolerance levels by age, gender, and participation in various activities were used to examine differences in tolerance of various levels of algae. Groups that demonstrated a lower tolerance for algae included older respondents, females, and those who had not participated in activities on or in West Virginia waters (RM 2012).

West Virginia DEP realized several positive outcomes from conducting the survey. Survey participants appreciated being asked their opinions about algae levels they found acceptable in surface waters. West Virginia DEP also found the survey results to be useful for communicating to the public about how the threshold for acceptable algae levels was derived. Additionally, the survey results provided West Virginia DEP with credible results for use in obtaining legislative and community buy-in and support to move forward with restoring surface waters with designated uses impacted by algae. West Virginia DEP has received only positive feedback about the user perception survey (Summers 2017, interview).

[1] Title 47 Code of State Rules (CSR) Series 2, *Requirements Governing Water Quality Standards*, West Virginia Department of Environmental Protection. https://apps.sos.wv.gov/adlaw/csr/.

Descriptive statistical analyses can be performed for all survey results to determine the distributional characteristics of the variables in the survey. These include calculations that determine a variable's central tendency (mean, median, mode) and dispersion (standard deviation, variance). It can also be useful to look at the frequency distribution and the shape of that distribution. These descriptive statistics can help determine what further types of statistical techniques are most appropriate for the data (Vaske 2008). Additional analyses, such as cross-tabulations, chi-square or t-tests, and different types of regression analyses can allow for further comparisons across variables and groups. Because categorical and ordinal data are common to surveys, special consideration should be given to the statistical methods used to analyze the various data collected.

When analyzing the survey results, the state should consider the temporal variability and timescales associated with the survey. As with discrete chemical or biological measurements, user perception measurements in survey results can vary over time as the water quality constituents (e.g., suspended sediment) and factors (e.g., seasonal changes in sunlight) that affect visual perception vary over time. Variation over time in the primary producer community's response to nutrients (i.e., uptake, cell division, and accumulation of biomass), which can occur on intra-annual and interannual timescales, can also contribute variability to visual perceptions. Users' perceptions and opinions about the aquatic resource could also change over time. For example, how users perceive water quality in a given waterbody could vary over time as a result of changes to water quality in other waterbodies that have shaped users' expectations. Changes in user demographics and the associated population's values could also change how a water is perceived. User perceptions could also simply differ over time because the water quality in the waterbody itself has changed over time, resulting in a revision of earlier expectations. Additionally, users could become more educated regarding water quality, causing their perception of a waterbody to change. Recognizing and accounting for these forms of variability during survey analysis can aid in characterizing and communicating survey results.

One interviewee recommended that states work closely with experts in survey research and statistics to determine the types of analyses required prior to designing the survey. Then, based on expert feedback, design the survey to collect the data needed for those analyses. Many state universities have statistics experts who may be able to provide assistance.

After analyzing the data and developing a target for user-perceived water quality for the protection of aesthetic and/or recreational designated uses, the state could then apply the target for developing numeric nutrient criteria. To do so, the state would analyze observed nutrient data and response data for the assessment endpoint and compare the result to the survey developed target. EPA has detailed in several guides the criteria development process using targets such as those developed using user perception surveys (USEPA 2000a, 2000b, 2001b).

A conceptual model for analyzing user perception survey results is provided in Figure 6.

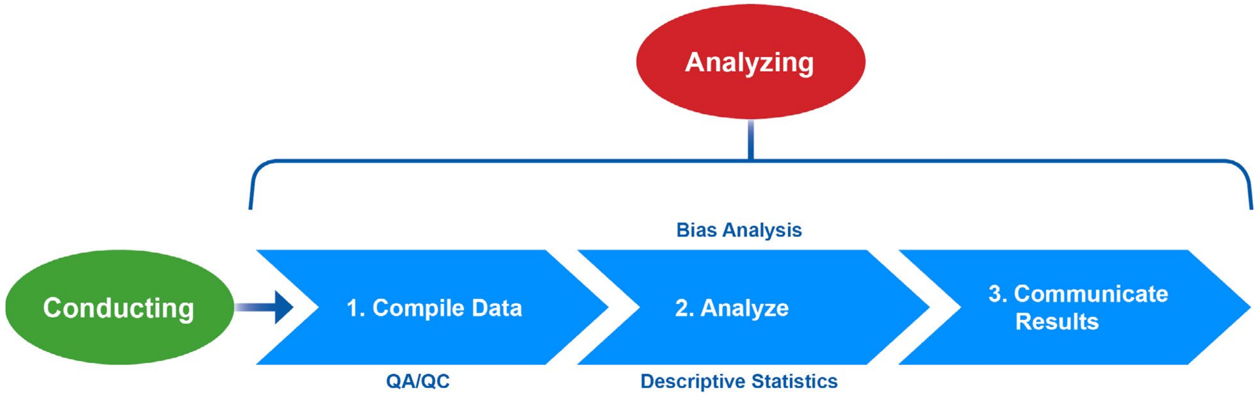

Figure 6. Conceptual model for analysis of user perception survey results

Shown in Figure 6 are steps a state could perform, as described in more detail below and in the section above, to analyze user perception survey results.

1. Compile data.
   Ensure QA/QC of data entry and analyses.
2. Analyze data.
   Identify sources of bias and adjust to minimize their influence.
   Develop descriptive statistics.
3. Present analysis and communicate results.

4.0 Ensuring Rigor in the Survey Process and Results

4.1 Quality Assurance/Quality Control

To ensure the quality of results generated from user perception surveys, a state should consider designing and implementing its survey in accordance with the following documents:

- EPA's *Survey Management Handbook* (USEPA 2003)
- A state QA project plan (QAPP) for survey development and implementation (if available)
- A state standard operating procedure (SOP) for conducting surveys (if available)

During the survey planning stage, it is useful for the state to define DQOs for the user perception survey on the basis of the resources available and project goals, and in accordance with EPA's technical planning process defined in *Guidance on Systematic Planning Using the Data Quality Objectives Process, EPA QA/G-4* (USEPA 2006). The DQO process includes identifying the quality and quantity of information required to make specific project decisions, regardless of scale. The DQO process includes defining not only the decisions, but also the tolerance for error in the decisions and, subsequently, the performance objectives for the data collection.

If a QAPP that covers survey design and development is not available for the state, the state will benefit from developing a QAPP in accordance with the specifications for format and content in *Requirements for Quality Assurance Project Plans*, *EPA QA/R-5* (USEPA 2001a) and *Guidance for Quality Assurance Project Plans, EPA QA/G-5* (USEPA 2002). In addition, it is recommended that a state develop an SOP for conducting surveys following EPA's *Guidance for Preparing Standard Operating Procedures (SOPs), EPA QA/G-6* (USEPA 2007); following the SOP will help ensure that surveys are conducted consistently and correctly.

QC has been addressed in several subsections of section 3.2 and 3.3, including survey error, sample size, pre-implementation testing, and picture selection. States should consider QC at all steps of survey design and implementation. An expert on survey methodology can help identify these considerations and support development of DQOs, a sampling plan for conducting surveys, and a QAPP for survey development and implementation.

Following are some examples of QC checks that should be employed during survey implementation and survey data processing, but this is by no means a complete list:

- Coordinate and control field work.
- Perform a nonresponse check to determine if nonrespondents are significantly different from those who responded to the survey.
- Review and edit questionnaires.
- Double-check tabulation calculations.

4.2 Maximizing Technical Rigor

Ensuring rigor in the survey process, data analysis, and results has been an important thread throughout this paper. Options have been identified that should increase the strength of a survey throughout the process, including hiring outside experts to help with the design and analysis, engaging stakeholder groups early, and targeting key user groups in the survey. The following highlights note important practices that could help ensure a high-quality survey process and product:

- If the state does not have sufficient in-house expertise, outsource all or part of the survey process to a third-party consultant with a good reputation and acknowledged expertise in the area. An academic institution could be considered.
- If the state decides to conduct the survey itself, consulting outside experts is still recommended to ensure that the most up-to-date and relevant techniques are used in the process.
- If conducting one-on-one interviews or leading a focus group, engage a neutral party to serve as the interviewer to avoid potential response bias.
- Seek IRB approval of survey methodology; an academic or research institution would know if involving an IRB is appropriate and, if so, which IRB the state should approach. An IRB can provide meaningful and useful input to help the survey process run smoothly. Sufficient time should be allotted to allow IRB review.
- Adopt a blended methodology, if possible, with a mix of on-site interviews, email, and mail.

As mentioned earlier, it is vitally important that the state and the survey team continually report and document the survey process from problem formulation through implementation and analysis, laying out the rationale for key decisions. This record provides a cross reference for QC and an administrative record that can help defend the survey process against challenges.

5.0 Survey Design Scenarios

Through the literature review and interviews used to inform this primer, three general scenarios—Scenarios 1, 2, and 3, below—and corresponding levels of effort were identified for completing a user perception survey for nutrient criteria development. Each of these three scenarios is suited for a particular set of circumstances that are determined by a state's needs, resources, and level of precision needed in the survey responses. These scenarios present general options, representing the range of activities that states could adopt. The scenarios, if properly applied and consistent with the intended survey objective(s), represent defensible approaches for developing user perception surveys for nutrient criteria development.

However, these scenarios are presented as general examples, recognizing that the set of circumstances in each survey can be unique. The is no single recommended approach for all states because each state has its own unique challenges and opportunities that will determine the best approach. It is recommended that states interested in surveys consider their specific situation to determine the approach that best meets their needs. Individuals experienced in survey design can help answer questions on how unique aspects of a state's situation affect their survey design. Detailed best practices information for each survey component are included in sections 3.0 and 4.0. This section provides a description of some key considerations that might affect how a state designs a user perception survey and a description of the three general scenarios.

5.1 Examples of Considerations when Designing a Survey

Below are examples of four considerations (existing data, existing program or stakeholder groups, funding, survey error) that might serve as a foundation for a state in designing a user perception survey. Each of the key criteria has been selected because it can create or identify existing constraints that limit the options available for conducting the survey. Each key consideration is elaborated on in the scenarios below.

5.1.1 Existing data

The existing data the state has for the waterbodies included in the potential survey.

Existing data can include current and historical water quality data, survey population information, and other chemical, physical, and biological data. Ideally, existing data sources used in designing a user perception survey will be peer-reviewed journal articles or assessments prepared by EPA, other federal entities (e.g., U.S. Geological Survey, National Oceanic and Atmospheric Administration), or state or local assessments (or assessments prepared under EPA or state guidance) for which project-specific QAPPs, sampling plans, or similar documentation describing site-specific sampling and analyses are available. Existing data can also include photographs with relevant metadata such as source, date, time, location, and associated nutrient sample data. Some data, such as population information and water quality data, are required to conduct *any* user perception survey of water quality, while other data (e.g., respondent age, sex, address, marital status, ethnicity, education, and income; how frequently a respondent visits a particular waterbody and purpose of visits [e.g., fishing, swimming]) open up opportunities for more robust survey approaches. The survey team should take into consideration any apparent inconsistencies or limitations of the existing data before using them.

5.1.2 Existing program or stakeholder groups

Information about the population, including population subsets of key stakeholders or previously engaged participants.

Population data help to determine the appropriate sample size and select a target response rate. Information on population subsets and key stakeholders is vitally important, particularly if one or more of the groups could potentially challenge the survey results or any resulting policy outcomes. Previously engaged participants might also be more perceptive than the general public and are not necessarily representative of the greater population.

This information also creates space for states who want to "go above and beyond" to realize indirect benefits to the survey process. Engaging key stakeholder and recreational user groups, in addition to the general public, early in the design process can develop and enhance buy-in to the process and the results, increase visibility of the process, improve response rates, and provide additional benefits.

5.1.3 Funding

The availability of funding to support survey design, implementation, and analysis.

The importance of funding availability in selecting and customizing a survey scenario cannot be downplayed. The three scenarios cover a range of price points and have different levels of staff resource needs.

5.1.4 Survey error

The consideration of error in survey design and how to best minimize total error within given time and resource constraints.

A state should determine what level of total survey error it is willing to accept when developing a survey. Generally, the lower the survey error rate, the greater the amount of resources—time, energy, and money—required to achieve it. Surveys can also be designed to directly address a particular survey error concern.

5.2 Scenario 1

This scenario is most suited for states with sufficient resources or that have unique issues such that leveraging similar surveys from nearby states is not practical or appropriate. The state also has funding to engage experts in survey methodology, statistical analysis, and stakeholder engagement.

5.2.1 Design

The survey is designed through a rigorous process. The team conducts preliminary work to identify specific survey needs and objectives, including response rate, target rate of statistical significance (e.g., 95 percent), and project timeline. In addition, the state engages key stakeholder groups early and often throughout the initiative.

5.2.2 Method

The survey is designed as a photo-based survey. Images have metadata on nutrient levels sampled at the time and location of the image. Photos are selected with appropriate variation in nutrient levels and consistency in waterbody presentation. The photos are tested with informed and lay audiences. Once the survey is designed, the state convenes focus groups of random members of the public to test out the survey questions and images. The survey team incorporates feedback, revises the survey, and retests it.

5.2.3 Delivery

The survey is delivered via mail, digitally, and in person. Using voter registration lists, surveys are distributed to a random sample of the population. Recipients have the option of responding with a mail-in form or by using an online form. Surveyors visit waterbodies to conduct in-person surveys with members of the public recreating on, in, or near the waterbodies. Key stakeholder groups are engaged to collaborate on survey distribution. The survey is conducted during a predetermined timeframe during the season(s) with the highest levels of recreational activity.

5.2.4 Analysis

Once the survey responses are collected, independent experts experienced in survey statistics analyze the results, comparing the responses within and across the delivery methods. Response analysis includes a detailed assessment of the survey respondents, identifying and adjusting for bias for each delivery method and across all survey responses.

5.2.5 Synopsis

This scenario requires more resources than the other scenarios. The careful design and testing process, including stakeholder engagement, allows the team to identify challenges early and incorporate them into survey design. The mixed-mode framework incorporating multiple survey delivery methods increases overall response rate, while providing more data to identify and adjust for potential bias. Independent analysis of survey results can be helpful to ensure unbiased interpretation of the results. The entire survey process purposefully works to minimize total survey error.

5.3 Scenario 2

This scenario involves a state that determines it would like to evaluate whether narrative nutrient criteria are being met and eventually develop aesthetic-based numeric nutrient criteria, but does not have an immediate need to develop those numeric nutrient criteria. While it has some resources, the state does not want to spend a lot of money or staff time designing and implementing a user perception survey. Instead, the state decides to leverage its ongoing volunteer monitoring and sampling program, using surveys to query volunteers' perceptions during their normal monitoring activities.

5.3.1 Design

The volunteer program has been in place for many years and has a group of dedicated volunteers who live on various waterbodies around the state and take water samples. It would be relatively easy to add a short survey to the samplers' routine out in the field. Such an approach to user perception surveys is very low cost and easy to implement, and all survey results are directly linked to water quality data taken at the same time as the survey. It is also a potential tool to collect data over years or even decades, which allows the state to track historic trends.

5.3.2 Method

The state drafts a survey inquiring about the respondents' perception of aesthetics of the waterbody and/or their desire to recreate in the waterbody. The state develops detailed instructions for how volunteers should respond in the course of their related activities. Pending resources, the state tests the detailed instructions and survey format with a small group of volunteers.

5.3.3 Delivery

The state distributes surveys to its volunteer group(s), including detailed instructions for completing the survey. Where appropriate, the state provides in-person briefings to volunteer groups on how to participate in the survey. The survey is carried out by the monitoring volunteers before they take their typical samples and measurements. The volunteers submit their answers to program authorities regularly as part of their annual volunteer program activities.

5.3.4 Analysis

After approximately 3 years of collecting the survey data, the state determines it has collected enough survey data to begin evaluating whether narrative nutrient criteria are being met and to support the development of numeric nutrient criteria. It chooses to analyze the results using in-house resources. The state chooses to collect additional years of data to potentially highlight volunteer transitions as the respondent group changes composition. Analysis can become complicated as staff correlate each user perception survey response with the water quality data captured by the respondents over time. The state uses its findings to determine whether narrative nutrient criteria are being met and to support the development of numeric nutrient criteria.

5.3.5 Synopsis

Scenario 2 carries a higher risk than Scenario 1 for coverage and nonresponse error since the state only collects survey data from a specific subgroup with a known interest in water quality. This subgroup may also have greater sensitivity to changes in water quality than the general public because of training and expertise related to their sampling practices.[15] The user perception data is directly correlated with water quality data, which could be useful for developing numeric nutrient criteria, but the volunteer perception data also might complicate analyses since these users are probably not representative of the general population. This scenario has a lower level of effort and presumably lower costs. It also allows ongoing information collection, providing an opportunity to analyze trends in perception over time. Such an approach is better for states that want to leverage their ongoing volunteer monitoring and sampling programs.

5.4 Scenario 3

In the third scenario, a state needs to set numeric nutrient criteria. The state, however, does not have the resources necessary to design *and* implement its own survey. Two nearby states have recently set the same numeric nutrient criteria based on intensive user perception survey processes that incorporated most or all of the aspects of the approach in Scenario 1. The state in question leverages the survey experience of the nearby states to streamline their survey process.

5.4.1 Design

The state determines how best to apply the other states' criteria to its own circumstances after examining its unique natural resource make-up (compared to the other states). The state also decides to use feedback from stakeholders to gauge their thoughts on applying the criteria in their state. Additionally, the state compares its own designated uses with those of the other states to ensure criteria are compatible.

5.4.2 Method

The state first determines the similarity of the waterbodies of interest and the communities surveyed in the other states to those in its own state. It then examines the survey design and assumptions to ensure they are relevant to the its own situation.

5.4.3 Delivery

If it finds that its own situation is similar to those of the other two states, the state adopts the same nutrient criteria as those states. Under this scenario, delivery of a survey would not be needed. However, sharing the survey through outreach to the public, and possibly particular stakeholder groups, to familiarize them with the details of the survey used in the other states is found to be helpful for buy-in.

5.4.4 Analysis

Although no direct analysis of data is needed, the state reviews the analyses done by the other states. This helps to ensure that the analyses done by those states address the same need and same question(s) as it is interested in.

[15] Additional information outlining state experiences with bias between pre-interested respondents and the general public is included in section 3.2 of this primer.

5.4.5 Synopsis

Scenario 3's approach has a higher risk for survey coverage error than Scenario 1, but also costs significantly less. It would be ideal to leverage the effort and resources already used by the other two states, as long as the designated uses, populations, and waterbodies of interest are similar enough. If the context of the survey carried out in another state is vastly different, the applicability of adopting the same numeric nutrient criteria decreases. In addition to taking on the total survey error that is present in the surveys conducted in the other states, this approach also carries significant risk of coverage errors if the populations in the states significantly differ. If the state adequately tests and accounts for this, however, the risks can be decreased.

The state could undertake additional actions to decrease the survey errors associated with this scenario. It could use a targeted approach to reach out to specific stakeholder or user groups to assess their perceptions. It could undertake outreach through either user groups or small-scale surveys to make sure that the proposed numeric nutrient criteria are in line with individual perceptions and to get buy-in for the criteria. The state would then internally determine how best to address and incorporate stakeholder comments and concerns (if any). If there are no major issues, the state has increased the robustness of fully adopting the same numeric nutrient criteria as its neighbors. The state can further fend off potential challenges by getting buy-in from key groups.

5.5 Summary of Survey Design Considerations

A summary of possible survey scenarios described in this primer with corresponding information on relative precision and resource needs is presented in Table 3.

Table 3. Summary of possible survey scenarios

Scenario	Precision	Resource Need
Volunteer sampling program	Low	Low
Survey of key user and stakeholder groups	Medium-High	Medium
Survey of user groups and a general population survey	High	High
General population survey	Medium	Medium-High
Copy another state	Medium-Low	Low
Copy another state and use targeted surveys	Medium-High	Low-Medium

Note: The scale of precision and resource needs is only a general relative estimate. The actual level will depend on the survey design and implementation plan adopted by the state.

Table 4 summarizes different options when conducting a survey, along with the relative cost, staff time, and notes on design considerations.

Table 4. Survey design option considerations

Survey Design Option Consideration	Cost	Staff Time	Notes
Expertise			
Engage outside experts to help design and implement the survey and analyze the results	$$$$	🕑	Increases robustness greatly
Contract with a stakeholder engagement expert to identify key user and stakeholder groups, and facilitate focus groups	$$	🕑	Increases robustness
Presurvey			
Engage with key stakeholders to get input and buy-in for the survey design and methodology	$	🕑🕑	Increases robustness
Pretest survey with focus groups	$$	🕑🕑	Increases robustness
Survey Options[16]			
Mail survey	$	🕑🕑	Low response rate; follow-up mailings are a good practice
On-site in-person surveys	$$	🕑🕑🕑	High response rate
Online survey to targeted user groups	$	🕑	
Survey of a volunteer sampling group	$	🕑	Nearly 100% response rate; high risk for coverage error
Mixed or blended approach	$-$$$$	🕑 - 🕑🕑🕑	Cost and staffing needs will depend on the ultimate survey mode(s) chosen

A user perception survey can be expensive, and states may not have the resources necessary to conduct a full survey using the best practices of survey design. It is a best practice for a state to conduct a smaller survey well rather than attempting to conduct a large survey poorly. In the case in which a state determines it does not have the resources to conduct a high-quality statewide user perception survey, it could consider possible measures to reduce resource needs such as those listed in section 3.3.1.7. Additionally, it could consider conducting a pilot survey of a small number of waterbodies. It could also lay additional groundwork for a full user perception survey in the future by collecting data and information that would be helpful for designing a survey, such as taking photographic images or determining popularity of specific waterbodies among visitors.

6.0 Conclusion

States have a variety of tools they can use when deriving numeric nutrient criteria. Among those options, user perception surveys provide a unique and flexible method to examine the effects of nutrient pollution on the aesthetic and recreational qualities of the nation's surface waters. In this role they are a useful option for a state to consider when it is determining the different ways to derive criteria.

Because the judgment of aesthetic quality is measured by the senses of the perceiver, user surveys allow states to directly measure whether a water is fulfilling its aesthetic uses and, by association, its

[16] The relative cost estimate of conducting different survey modes assumes that the state has access to the relevant populations and groups and their contact information to be able to conduct the survey effectively. Acquiring this information might require additional resources and staff time.

recreational uses. By directly tying user experiences to the criteria being developed, the state creates credibility both for itself and the criteria.

User perception surveys can also help states directly engage the public, creating public awareness of the water resources. Members of an engaged and aware public often feel empowered to become active citizens and encouraged to participate in protecting water resources, thus strengthening the state. Public participation in locally driven, bottom-up management can be an influential factor in the long-term success of water quality protection. Public interest and involvement can be particularly advantageous in states where local communities are authorized to make decisions regarding water quality protection or where nonpoint source nutrient contributions are prevalent, often requiring voluntary measures or government incentives or policies that need public support.

Given the positive benefits a user perception survey can provide in developing numeric nutrient criteria, it is an important tool with which a state can address nutrient pollution. States can use the information presented in this primer to determine if a survey is the appropriate tool to help achieve their goals and, if so, familiarize themselves with the technical aspects and decisions needed to design an effective user perception survey. Additional resources for a more detailed examination of the intricate aspects of user perception surveys can be found in the literature used to develop this paper.

Appendix A. References

ADEQ (Arizona Department of Environmental Quality). 2009. *Title 18. Environmental Quality: Chapter 11. Department of Environmental Quality Water Quality Standards: Article 11. Water Quality Standards for Surface Waters*. Arizona Department of Environmental Quality, Phoenix, AZ. Accessed May 2017. http://legacy.azdeq.gov/environ/water/standards/download/SWQ_Standards-1-09-unofficial.pdf.

Backer, L.C., L. Fleming, A. Rowan, and D. Baden. 2003. Epidemiology and Public Health of Human Illnesses Associated with Harmful Marine Algae. Chapter 26 in *Manual on Harmful Marine Microalgae*, ed. G.M. Hallegraeff, D.M. Anderson, and A.D. Cembella, pp. 723–749. UNESCO Publishing, Paris. http://unesdoc.unesco.org/images/0013/001317/131711e.pdf.

Bricker, S.B., C.G. Clement, D.E. Pirhalla, S.P. Orlando, and D.R.G. Farrow. 1999. *National Estuarine Eutrophication Assessment: Effects of Nutrient Enrichment in the Nation's Estuaries*. National Oceanic and Atmospheric Administration, National Ocean Service, Special Projects Office and the National Centers for Coastal Ocean Science, Silver Spring, MD.

Brown, T.C., and T.C. Daniel. 1991. Landscape aesthetics of riparian environments: Relationship of flow quantity to scenic quality along a wild and scenic river. *Water Resources Research* 27(8):1787–1795.

CDEEP (Connecticut Department of Energy and Environmental Protection). 2013. *Water Quality Standards*. State of Connecticut Department of Energy and Environmental Protection. Accessed May 2017. https://portal.ct.gov/DEEP/Water/Water-Quality/Water-Quality-Standards-and-Classification.

CEPA (California Environmental Protection Agency). 2015. *California Ocean Plan, Water Quality Control Plan, Ocean Waters of California*. California Environmental Protection Agency, State Water Resources Control Board, Sacramento, CA. Accessed March 2017. http://www.waterboards.ca.gov/water_issues/programs/ocean/docs/cop2015.pdf.

Dahlhamer, J.M., M.L. Cynamon, J.F. Gentleman, A. Piani, and M.J. Weiler. 2010. Minimizing Survey Error Through Interviewer Training: New Procedures Applied to the National Health Interview Survey. In *Proceedings of the Joint Statistical Meetings*, American Statistical Association, Vancouver, British Columbia, July 31–August 5, 2010, pp. 4627–4640. Accessed March 2017. http://www.asasrms.org/Proceedings/y2010/Files/308678_61376.pdf.

Daniel, T.C., and J. Vining. 1983. Methodological Issues in the Assessment of Landscape Quality. Chapter 2 in *Behavior and the Natural Environment,* ed. I. Altman and J.F. Wohlwill, pp. 39–84. Plenum Press, New York.

Dillman, D.A., J.D. Smyth, and L.M. Christian. 2014. *Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method*. 4th ed. John Wiley & Sons, Inc., Hoboken, New Jersey.

Ditton, R.B., and T.L. Goodale. 1973. Water quality perception and the recreational uses of Green Bay, Lake Michigan. *Water Resources Research* 9(3):569–579.

Dodds, W.K., W.W. Bouska, J.L. Eitzmann, T.J. Pilger, K.L. Pitts, A.J. Riley, J.T. Schloesser, and D.J. Thornbrugh. 2008. Eutrophication of U.S. freshwaters: Analysis of potential economic damages. *Environmental Science and Technology* 43(1):12–19.

Dodds, W.K., and V.H. Smith. 2016. Nitrogen, phosphorus, and eutrophication in streams. *Inland Waters* 6(2):155–164.

Egan, K.J., J.A. Herriges, C.L. King, and J.A. Downing. 2009. Valuing water quality as a function of water quality measures. *American Journal of Agricultural Economics* 91(1):106–123.

Glibert, P.M., C.J. Madden, W. Boynton, D. Flemer, C. Heil, and J. Sharp. 2010. *Nutrients in Estuaries: A Summary Report of the National Estuarine Experts Workgroup 2005–2007*. U.S. Environmental Protection Agency, Office of Water, Washington, DC. Accessed March 2017. https://www.epa.gov/sites/production/files/documents/nutrients-in-estuaries-november-2010.pdf.

Heiskary, Steven, Minnesota Pollution Control Agency [retired]. 2017, September 6. Telephone interview with Erica Wales, Kearns & West, regarding a case study for EPA's user perception survey white paper.

Heiskary, S.A., and W.W. Walker, Jr. 1988. Developing phosphorus criteria for Minnesota lakes. *Lake and Reservoir Management* 4(1):1–9.

Heiskary, S.A., and W.W. Walker, Jr. 1995. Establishing a chlorophyll *a* goal for a run-of-the-river reservoir. *Lake and Reservoir Management* 11(1):67–76.

Heiskary, S.A., and C.B. Wilson. 2005. *Minnesota Lake Water Quality Assessment Report: Developing Nutrient Criteria*. 3rd ed. Minnesota Pollution Control Agency, Saint Paul, MN. Accessed March 2017. http://www.pca.state.mn.us/index.php/view-document.html?gid=6503.

Heisler, J., P.M. Glibert, J.M. Burkholder, D.M. Anderson, W. Cochlan, W.C. Dennison, Q. Dortch, C.J. Gobler, C.A. Heil, E. Humphries, A. Lewitus, R. Magnien, H.G. Marshall, K. Sellner, D.A. Stockwell, D.K. Stoecker, and M. Suddleson. 2008. Eutrophication and harmful algal blooms: A scientific consensus. *Harmful Algae* 8(1):3–13.

Hetherington, J., T.C. Daniel, and T.C. Brown. 1993. Is motion more important than it sounds? The medium of presentation in environment perception research. *Journal of Environmental Psychology* 13(4):283–291.

Hilborn, E.D., V.A. Roberts, L. Backer, E. Deconno, J.S. Egan, J.B. Hyde, D.C. Nicholas, E.J. Wiegert, L.M. Billing, M. Diorio, M.C. Mohr, J.F. Hardy, T.J. Wade, J.S. Yoder, and M.C. Hlavsa. 2014. Algal bloom-associated disease outbreaks among users of freshwater lakes—United States, 2009–2010. *Centers for Disease Control and Prevention, Morbidity and Mortality Weekly Report* 63(1):11–15.

House, M.A., and E.K. Sangster. 1991. Public perception of river-corridor management. *Water and Environment Journal* 5(3):312–316.

Hoyer, M.V., C.D. Brown, and D.E. Canfield, Jr. 2004. Relations between water chemistry and water quality as defined by lake users in Florida. *Lake and Reservoir Management* 20(3):240–248.

KDEP (Kentucky Department for Environmental Protection). 2013. *401 KAR 10.031. Surface water standards*. Department for Environmental Protection, Division of Water, Frankfort, KY. Accessed May 2017. http://www.lrc.ky.gov/kar/401/010/031.htm.

Keeler, B.L., S.A. Wood, S. Polansky, C. Kling, C.T. Filstrup, and J.A. Downing. 2015. Recreational demand for clean water: Evidence from geotagged photographs by visitors to lakes. *Frontiers in Ecology and the Environment* 13(2):76–81.

Kishbaugh, S.A. 1994. Applications and limitations of qualitative lake assessment data. *Lake and Reservoir Management* 9(1):17–23.

Kohavi, R., R. Longbotham, D. Sommerfield, and R.M. Henne. 2009. Controlled experiments on the web: Survey and practical guide. *Data Mining and Knowledge Discovery* 18:140–181. http://ai.stanford.edu/~ronnyk/2009controlledExperimentsOnTheWebSurvey.pdf.

Lopez, C.B., E.B. Jewett, Q. Dortch, B.T. Walton, and H.K. Hudnell. 2008. *Scientific Assessment of Freshwater Harmful Algal Blooms*. Interagency Working Group on Harmful Algal Blooms, Hypoxia, and Human Health of the Joint Subcommittee on Ocean Science and Technology, Washington, DC.

MDEQ (Michigan Department of Environmental Quality). 2006. *Part 4. Water Quality Standards*. Michigan Department of Environmental Quality, Water Bureau, Lansing, MI. Accessed March 2017. http://www.michigan.gov/documents/deq/wrd-rules-part4_521508_7.pdf.

NAS and NAE (National Academy of Sciences and National Academy of Engineering). 1972. *Water Quality Criteria 1972: A Report of the Committee on Water Quality Criteria*. EPA-R3-73-033. Prepared for U.S. Environmental Protection Agency by Environmental Sciences Board, National Academy of Sciences and National Academy of Engineering, Washington, DC. Accessed March 2017. http://nepis.epa.gov/Exe/ZyPDF.cgi/2000XOYT.PDF?Dockey=2000XOYT.PDF.

Nicolson, J.A., and A.C. Mace, Jr. 1975. Water quality perception by users: Can it supplement objective water quality measures? *Water Resources Bulletin* 11(6):1197–1207.

NRC (National Research Council). 2000. *Clean Coastal Waters: Understanding and Reducing the Effects of Nutrient Pollution*. The National Academies Press, Washington, DC.

NYSDEC (New York State Department of Environmental Conservation). n.d. Citizens Statewide Lake Assessment Program (CSLAP). Page on NYSDEC website. Accessed March 2017. http://www.dec.ny.gov/chemical/81576.html.

NYSFOLA and NYSDEC (New York State Federation of Lake Associations and New York State Department of Environmental Conservation). 2003. *Evaluating Lake Perception Data as a Means to Identify Reference Nutrient Conditions.* Final Report to the U.S. Environmental Protection Agency Regions I, II, and V.

RM (Responsive Management). 2012. *West Virginia Residents' Opinions on and Tolerance Levels of Algae in West Virginia Waters*. Conducted for the West Virginia Department of Environmental Protection. Accessed March 2017. http://www.dep.wv.gov/WWE/Programs/wqs/Documents/WVAlgaeSurveReport_ResMgmt_WVDEP_2012.pdf.

RIDEM (Rhode Island Department of Environmental Management). 2009. *Water Quality Regulations*. Rhode Island Department of Environmental Management, Office of Water Resources, Providence, RI. Accessed March 2017. http://www.dem.ri.gov/pubs/regs/regs/water/h20q09a.pdf.

Smeltzer, Eric, Vermont Department of Environmental Conservation [retired]. 2017, September 6. Telephone interview with Erica Wales, Kearns & West, regarding a case study for EPA's user perception survey white paper.

Smeltzer, E., and S.A. Heiskary. 1990. Analysis and applications of lake user survey data. *Lake and Reservoir Management* 6(1):109–118.

Smeltzer, E., N.C. Kamman, and S. Fiske. 2016. Deriving nutrient criteria to minimize false positive and false negative water use impairment determinations. *Lake and Reservoir Management* 32(2):182–193.

Smith, A.J., B.T. Duffy, and M.A. Novak. 2015. Observer rating of recreational use in wadeable streams of New York State, USA: Implications for nutrient criteria development. *Water Research* 69:195–209.

Smith, D.G., A.M. Cragg, and G.F. Croker. 1991. Water clarity criteria for bathing waters based on user perception. *Journal of Environmental Management* 33(3):285–299.

Smith, D.G., and R.J. Davies-Colley. 1992. Perception of water clarity and colour in terms of suitability for recreational use. *Journal of Environmental Management* 36(3):225–235.

Smith, D.G., G.F. Croker, and K. McFarlane. 1995a. Human perception of water appearance 1. Clarity and colour for bathing and aesthetics. *New Zealand Journal of Marine and Freshwater Research* 29(1):29–43.

Smith, D.G., G.F. Croker, and K. McFarlane. 1995b. Human perception of water appearance 2. Colour judgment, and the influence of perceptual set on perceived water suitability for use. *New Zealand Journal of Marine and Freshwater Research* 29(1):45–50.

Smith, V.H., G.D. Tilman, and J.C. Nekola. 1999. Eutrophication: Impacts of excess nutrient inputs on freshwater, marine, and terrestrial ecosystems. *Environmental Pollution* 100(1–3):179–196.

Smith, V.H. 2003. Eutrophication of freshwater and coastal marine ecosystems: A global problem. *Environmental Science and Pollution Research* 10(2):126–139.

Summers, James, West Virginia Department of Environmental Protection. 2017, September 29. Telephone interview with Erica Wales, Kearns & West, regarding a case study for EPA's user perception survey white paper.

Suplee, M., V. Watson, A. Varghese, and J. Cleland. 2008. *Scientific and Technical Basis of the Numeric Nutrient Criteria for Montana's Wadeable Streams and Rivers*. Montana Department of Environmental Quality, Helena, MT.

Suplee, M.W., V. Watson, M. Teply, and H. McKee. 2009. How green is too green? Public opinion of what constitutes undesirable algae levels in streams. *Journal of the American Water Resources Association* 45(1):123–40.

Suplee, Michael, Montana Department of Environmental Quality. 2017, September 1. Telephone interview with Erica Wales, Kearns & West, regarding case study for EPA's user perception survey white paper.

TCEQ (Texas Commission on Environmental Quality). 2014. Texas Surface Water Quality Standards. Texas Commission on Environmental Quality, Austin, TX. http://texreg.sos.state.tx.us/public/readtac$ext.TacPage?sl=R&app=9&p_dir=&p_rloc=&p_tloc=&p_ploc=&pg=1&p_tac=&ti=30&pt=1&ch=307&rl=4.

TWCA (Texas Water Conservation Association). 2005. *Development of Use-Based Chlorophyll Criteria for Recreational Uses of Reservoirs.* Final report. Presented to Texas Commission on Environmental Quality. Prepared by Brazos River Authority, Guadalupe-Blanco River Authority, Lower Colorado River Authority, Sabine River Authority, San Antonio River Authority, Tarrant Regional Water District, and Trinity River Authority.

USEPA (U.S. Environmental Protection Agency). 1998. *Guidelines for Ecological Risk Assessment*. EPA/630/R-95/002F. U.S. Environmental Protection Agency, Washington, DC. Accessed July 2017. https://nepis.epa.gov/Exe/ZyPDF.cgi/30004XFR.PDF?Dockey=30004XFR.PDF.

USEPA (U.S. Environmental Protection Agency). 2000a. *Nutrient Criteria Technical Guidance Manual: Rivers and Streams*. EPA/822/B-00/002. U.S. Environmental Protection Agency, Office of Water, Washington, DC. Accessed March 2017. https://nepis.epa.gov/Exe/ZyPDF.cgi/20003CVP.PDF?Dockey=20003CVP.PDF.

USEPA (U.S. Environmental Protection Agency). 2000b. *Nutrient Criteria Technical Guidance Manual: Lakes and Reservoirs*. EPA/822/B-00/001. U.S. Environmental Protection Agency, Office of Water, Washington, DC. Accessed March 2017. https://nepis.epa.gov/Exe/ZyPDF.cgi/20003COV.PDF?Dockey=20003COV.PDF.

USEPA (U.S. Environmental Protection Agency). 2001a. *Requirements for Quality Assurance Project Plans, EPA QA/R-5*. EPA/240/B-01/003. U.S. Environmental Protection Agency, Office of Environmental Information, Washington, DC. Accessed March 2017. https://www.epa.gov/sites/production/files/2016-06/documents/r5-final_0.pdf.

USEPA (U.S. Environmental Protection Agency). 2001b. *Nutrient Criteria Technical Guidance Manual: Estuarine and Coastal Marine Waters*. EPA/822/B-01/003. U.S. Environmental Protection Agency, Office of Water, Washington, DC. Accessed March 2017. https://nepis.epa.gov/Exe/ZyPDF.cgi/20003FDF.PDF?Dockey=20003FDF.PDF.

USEPA (U.S. Environmental Protection Agency). 2002. *Guidance for Quality Assurance Project Plans, EPA QA/G-5*. EPA/240/R-02/009. U.S. Environmental Protection Agency, Office of Environmental Information, Washington, DC. Accessed March 2017. https://www.epa.gov/sites/production/files/2015-06/documents/g5-final.pdf.

USEPA (U.S. Environmental Protection Agency). 2003. *Survey Management Handbook*. EPA 260-B-03-003. U.S. Environmental Protection Agency, Office of Information Analysis and Access, Washington, DC. Accessed March 2017. https://nepis.epa.gov/Exe/ZyPDF.cgi/P1005GNB.PDF?Dockey=P1005GNB.PDF.

USEPA (U.S. Environmental Protection Agency). 2006. *Guidance on Systematic Planning Using the Data Quality Objectives Process, EPA QA/G-4.* EPA/240/B-06/001. U.S. Environmental Protection Agency, Office of Environmental Information, Washington, DC. Accessed March 2017. https://www.epa.gov/sites/production/files/2015-06/documents/g4-final.pdf.

USEPA (U.S. Environmental Protection Agency). 2007. *Guidance for Preparing Standard Operating Procedures, EPA QA/G-6.* EPA/600/B-07/001. U.S. Environmental Protection Agency, Office of Environmental Information, Washington, DC. Accessed March 2017. https://www.epa.gov/sites/production/files/2015-06/documents/g6-final.pdf.

USEPA (U.S. Environmental Protection Agency). 2016. *Assessment and Total Maximum Daily Load Tracking and Implementation System (ATTAINS)*. U.S. Environmental Protection Agency. Accessed March 2017. https://www.epa.gov/waterdata/assessment-and-total-maximum-daily-load-tracking-and-implementation-system-attains.

UDEQ (Utah Department of Environmental Quality). 2011. *Utah's Lakes & Rivers Recreation Survey 2011* (Survey form). Utah Department of Environmental Quality, Division of Water Quality, Salt Lake City.

Vaske, J.J. 2008. *Survey Research and Analysis: Applications in Parks, Recreation and Human Dimensions*. Venture Publishing, State College, PA.

VDEC (Vermont Department of Environmental Conservation). 2016. *Nutrient Criteria for Vermont's Inland Lakes and Wadeable Streams.* Technical Support Document, Revision. Vermont Department of Environmental Conservation, Watershed Management Division, Montpelier, VT. Accessed March 2017. http://dec.vermont.gov/sites/dec/files/wsm/Laws-Regulations-Rules/2016_12_22-Nutrient_criteria_technical_support_document.pdf.

Vollenweider, R.A. 1968. *Scientific Fundamentals of the Eutrophication of Lakes and Flowing Waters with Particular Reference to Nitrogen and Phosphorus as Factors in Eutrophication.* Technical Report DAS/CS1/68.27. Organization for Economic Co-operation and Development, Directorate for Scientific Affairs, Paris.

Welch, E.B., J.M. Jacoby, R.R. Horner, and M.R. Seeley. 1988. Nuisance biomass levels of periphytic algae in streams. *Hydrobiologia* 157:161–168.

West, A.O., J.M. Nolan, and J.T. Scott. 2016. Optical water quality and human perceptions of rivers: An ethnohydrology study. *Ecosystem Health and Sustainability* 2(8):1–11.

WHO (World Health Organization). 1999. *Toxic Cyanobacteria in Water: A Guide to Their Public Health Consequences, Monitoring and Management*, ed. I. Chorus and J. Bartram. E & EF Spon, London, England.

WHO (World Health Organization). 2003. *Guidelines for Safe Recreational Water Environments: Volume 1, Coastal and Fresh Waters*. World Health Organization, Geneva, Switzerland. Accessed March 2017. http://apps.who.int/iris/bitstream/10665/42591/1/9241545801.pdf.

WV DEP (West Virginia Department of Environmental Protection). 2017. Filamentous Algae in West Virginia. Page on WV DEP website. West Virginia Department of Environmental Protection, Charleston, WV. https://dep.wv.gov/WWE/watershed/Algae/Pages/-Nutrients-and-Filamentous-Algae-in-West-Virginia.aspx.

Appendix B. Interviews

Heiskary, Steven. Minnesota Pollution Control Agency. January 21, 2015.

Kishbaugh, Scott. New York State Department of Environmental Conservation. January 28, 2015.

Laidlaw, Tina. U.S. EPA Region 8. January 30, 2015.

Ostermiller, Jeff. Utah Department of Environmental Quality. February 18, 2015.

Smeltzer, Eric. Vermont Department of Environmental Conservation. January 21, 2015.

Summers, James. West Virginia Department of Environmental Protection. September 9, 2017.

Suplee, Mike. Montana Department of Environmental Quality. February 6, 2015.

Walker, Jr., William W. Consultant. January 29, 2015.

Appendix C. Survey Design Checklist and Questionnaire

Problem Formulation

1. Do changes in nutrient concentrations in a waterbody cause responses that can be visually observed?
2. Are these visual changes consistent among the waterbodies in question (e.g., do all of the streams of interest respond similarly)?
3. Are recreational users able to detect gradations of these visual changes?
4. Are these visual changes meaningful to recreational users or tribes?

Scoping

1. What are the criteria and conditions of the waterbodies of interest?
2. What are your key stakeholder and user groups?
   What resources are available to help you contact them?
3. What level of financial resources do you have to conduct the survey?
4. What types of water quality data are available? What resources and information are available to support survey development?
5. Do you have enough staff time to conduct the survey?
6. Do your in-house staff have expertise in:
   - Survey design?
   - Survey research and statistics?
7. What geographic scale will be used to implement the survey?
8. Is it necessary or possible to classify the waterbodies?
9. How do the answers to the scoping questions affect the options available for designing, conducting, and analyzing a survey? As the survey progresses, decisions should be made in light of the scoping questions.
10. Based on the decisions made in the previous steps, what type and amount of resources are needed?

Designing

1. What are your DQOs?
2. Do you plan to engage stakeholders? If so, how?
3. What population do you want represented in your survey population (e.g., general population, key user/stakeholder groups, both)?
   What key groups do you want to target? This assumes that the state has the ability to access those groups.
   Do you have access to contact information for a general population survey?
   Do you have access to contact information for key user or stakeholder groups?
4. What is your ideal sample size?
5. What steps will you take to minimize survey error?
   What margin of error is acceptable for your survey?

Conducting

1. What mode(s) will you use to conduct the survey?
   Looking at the comparison of survey modes, do you have adequate funding and staff resources to carry out the different modes you wish to use?
2. How will you select and refine questions and/or pictures used in the survey?
3. What are your plans for communicating with the public during and after the survey?
4. What types of information do you plan to collect about the demographics of the respondents?
5. How will you pretest or pilot your survey?
6. Do you plan to follow up with respondents after the survey and, if so, how?

Analyzing

1. How will you compile data?
   What methods will you use to ensure QA/QC of data entry and analyses?
2. What are potential analyses you plan to perform on the data?
   How will you identify sources of bias and adjust to minimize their influence?
   What descriptive statistics best characterize the data?
3. How, when, and to whom do you plan to present analyses and communicate results?