Fisheries and Oceans Canada / Pêches et Océans Canada
Ecosystems and Oceans Science / Sciences des écosystèmes et des océans

Canadian Science Advisory Secretariat (CSAS)

Research Document 2019/057

National Capital Region

A review of functional monitoring methods to assess mitigation, restoration, and offsetting activities in Canada

Douglas C. Braun[1], Karen E. Smokorowski[2], Michael J. Bradford[1], Luc Glover[2]

[1]Fisheries and Oceans Canada and School of Resource and Environmental Management
Simon Fraser University
8888 University Drive
Burnaby, BC V5A1S6

[2]Great Lakes Laboratory for Fisheries and Aquatic Sciences, Fisheries and Oceans Canada
1219 Queen Street East
Sault Ste. Marie, ON P6A 2E5

December 2019

Canada

Foreword

This series documents the scientific basis for the evaluation of aquatic resources and ecosystems in Canada. As such, it addresses the issues of the day in the time frames required and the documents it contains are not intended as definitive statements on the subjects addressed but rather as progress reports on ongoing investigations.

Published by:

Fisheries and Oceans Canada
Canadian Science Advisory Secretariat
200 Kent Street
Ottawa ON K1A 0E6

http://www.dfo-mpo.gc.ca/csas-sccs/
csas-sccs@dfo-mpo.gc.ca

ISSN 1919-5044

Correct citation for this publication:

Braun, D.C., Smokorowski, K.E., Bradford, M.J., and Glover, L. 2019. A review of functional monitoring methods to assess mitigation, restoration, and offsetting activities in Canada. DFO Can. Sci. Advis. Sec. Res. Doc. 2019/057. vii + 75 p.

Aussi disponible en français :

Braun, D.C., Smokorowski, K.E., Bradford, M.J., et Glover, L. 2019. Examen de la surveillance fonctionnelle pour évaluer les activités d'atténuation, de restauration et de compensation au Canada. Secr. can. de consult. sci. du MPO, Doc. de rech. 2019/057. vii + 81 p.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Fisheries and Oceans Canada’s Fisheries Protection Program (FPP) manages the impacts of habitat degradation or loss of fish productivity through management measures that include mitigation, restoration, and offsetting. To determine the effectiveness of these management measures, three levels of standardized monitoring have been recommended: 1) compliance, 2) functional, and 3) effectiveness monitoring. This report focuses on functional monitoring and provides a review of functional assessment approaches, methods, and the information needed to develop a standardized functional monitoring program. *Functional monitoring* is defined as a science-based, scaled-down version of effectiveness monitoring that relies on surrogate metrics to assess whether management measures provide expected conditions suitable for fish to carry out their life processes. Functional monitoring will use a quantitative approach that employs quick and easy to measure indicators of fish productivity. It will be applied to projects predicted to have low and/or well understood impacts. Objectives of a functional monitoring program are to determine if management measures recommended in the letter of advice, or required by the *Fisheries Act* authorization (e.g., mitigation, restoration, offsetting) are functioning as intended (Table 1). This report is the first step in providing science advice on the approaches, methods, and information required to develop a standardized functional monitoring program.

We conducted a literature review to gather information about how and where functional monitoring is conducted, and its key components. Important considerations for designing a functional monitoring program are identified and discussed in the following sections: 1) Monitoring Designs, 2) Rapid Assessments, 3) Metrics to Measure Function, and 4) Standardized Monitoring. Each section describes the benefits and challenges of common approaches and methods as well as considerations relevant to implementation.

The majority of studies we reviewed used the Reference Condition Approach, while other monitoring designs, mainly the Before-After-Control-Impact design, were used for longer-term monitoring of fish productivity or other biotic indicators, not habitat function. Rapid assessments were used in many of the studies reviewed, but definitions of what constituted a rapid assessment varied greatly. Clearly defining the level of effort to be invested in a functional monitoring program will be important in guiding the development of an effective program. We found functional assessments used a range of indicators, and selection of indicators for standardized functional monitoring may vary by region, system type, species, and life stage. There should be national oversight that ensures the results from monitoring projects are comparable among these organizational levels (e.g., region, system type, species, and life stage). We outline an example of a checklist approach for assigning indicators to project specific monitoring protocols, which is based on the FPP’s Pathways of Effects models. This approach allows for consistent use of indicators for projects where the same pathways of effects have been identified. Finally, the discussion of standardized functional monitoring programs highlights the value of such programs, but also the challenges and limitations.

INTRODUCTION

Fisheries and Oceans Canada's Fisheries Protection Program (FPP) manages impacts on fish productivity related to habitat degradation or loss, alterations to fish passage and flow, and aquatic invasive species (DFO 2018a). For projects that are expected to have impacts on fish productivity, the FPP administers letters of advice or *Fisheries Act* authorizations that provide guidance on how to avoid or mitigate impacts where possible, and requirements for restoration and offsetting where impacts are unavoidable and cannot be mitigated. The FPP reviews project results through monitoring to achieve two management monitoring objectives: 1) to ensure conformity with advice and compliance with the *Fisheries Act* and *Species at Risk Act* (i.e., did the proponent comply with the mitigation and/or compensation plan?); and 2) to evaluate the effectiveness of management measures (e.g., mitigation, restoration, and offsetting) aimed at reducing the impacts of projects on fish and fish habitat (i.e., was the habitat protected?) (Lewis et al. 2013). Monitoring projects provides the FPP with feedback information that can be used to adaptively manage at both the individual project and program levels.

Successful monitoring programs are based on well-defined questions, a conceptual understanding of relevant ecological processes and a robust study design that allows for inferences to be made about ecosystem change (Lindenmayer and Likens 2010) while also remaining adaptive to new information and questions (Lindenmayer and Likens 2009). Hierarchical monitoring strategies offer a range of monitoring approaches from simple routine assessments of compliance to estimates of change to environmental state to more complex assessments of a project's biological effectiveness and ecosystem function (Pearson et al. 2005). Importantly, monitoring programs are more likely to succeed if their purpose and objectives are defined by thoughtful scientifically-based questions (Lindenmayer and Likens 2010).

Hierarchical monitoring frameworks are a common approach for monitoring impacts (Hewitt et al. 2003) and the effectiveness of compensation activities (e.g., Koning et al. 1998, Schiff et al. 2011, DFO 2012). Once monitoring objectives (e.g., assessing state or detecting change) and the program's purpose have been established a hierarchical range of monitoring, strategies, designs, variables, methods etc. are available to identify an approach that will provide for efficient diagnostic power (Vos et al. 2000). Projects that are small with well understood impacts (e.g., installation of a culvert in a stream during road building) do not require a decade of monitoring to determine their effectiveness. A monitoring program similar to Pearson et al.'s (2005) routine monitoring may be sufficient to determine that the culvert was correctly installed and that the culvert is passing fish. Conversely, routine monitoring would not be sufficient for a project with potentially large impacts that are not well understood (e.g., stream diversion that will alter habitat for a species at risk). Pearson et al.'s (2005) and Smokorowski et al.'s (2015) have proposed a more complex effectiveness monitoring approach that has received national acceptance as a means to provide advice to FPP. However, there remains a need to develop scientifically defensible monitoring standards to assess biological function when smaller, well-understood impacts are anticipated.

PURPOSE OF DOCUMENT

The Fisheries Protection Program has requested advice from Science to ensure a more rigorous and systematic approach to functional monitoring when the FPP conducts monitoring and/or reviews monitoring results of low impact projects.

In a recent review, DFO (2012) identified three hierarchical levels of monitoring that have been recommended: 1) compliance, 2) functional, and 3) effectiveness monitoring (Figure 1).

Compliance monitoring is described as an operational monitoring activity used to evaluate whether the proponent has followed the advice or authorization issued by the FPP (DFO 2012). It is an audit of the construction standards that generally reflect best management practices (e.g., was the culvert built to the conditions outlined in the *Fisheries Act* authorization? Does the quantity of habitat built match the stated amount in the authorization?). Functional and effectiveness monitoring approaches are science-based activities that take further steps in evaluating projects in that their shared management monitoring objective is to determine if impacts on fish and fish habitat associated with a project have been effectively addressed through mitigation, restoration, or offsetting. These monitoring programs could be used in a tiered approach where they represent different starting points for meeting the same management objective. The order in which the monitoring programs are implemented might depend on the level of impact and the uncertainty in the outcomes of the management measures. For example, for projects that require more rigorous investigation (e.g., high impact and/or uncertain outcomes), effectiveness monitoring may be conducted at the beginning of the project, then transition to functional monitoring; while other projects that have lower impacts and/or well-understood outcomes may begin with functional monitoring, but continue with effectiveness monitoring if the initial results indicate a more in-depth investigation is warranted.

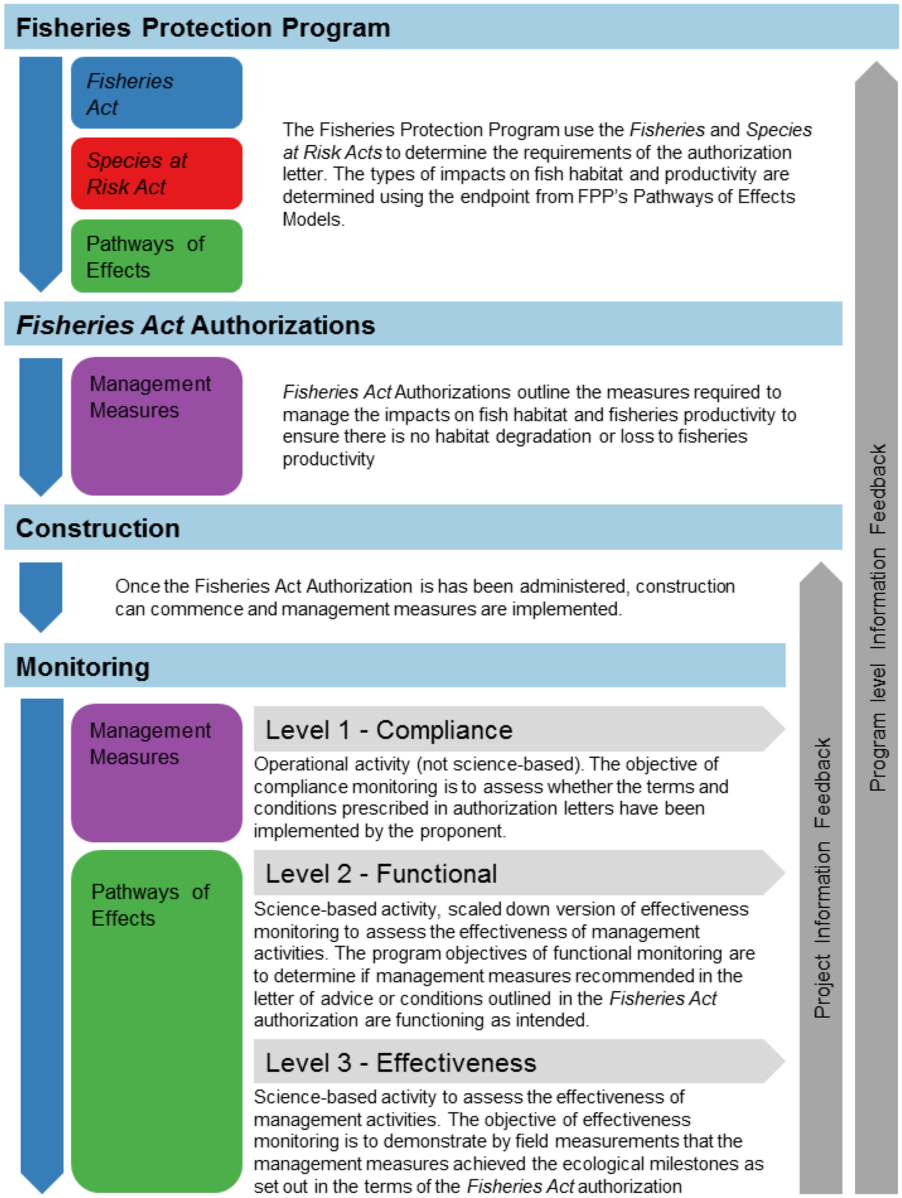


Figure 1. Conceptual diagram for monitoring the compliance and effectiveness of management measures outlined by authorizations by the Fisheries Protection Program (for illustrative purposes only). Coloured boxes indicate information or regulatory nodes that are used to inform each step of FPP's process, from developing Fisheries Act authorizations and management measures to mitigate project impacts, to monitoring the compliance and effectiveness of those measures. Workflow moves down the figure. Monitoring results can provide feedback information at the project and program levels in an adaptive management framework (dark grey vertical bars). Monitoring may inform the need for additional management or monitoring measures to be implemented by the proponent (construction stage), as well as the Fisheries Protection Program, and what management measures are most effective.

Based on the questions and hypotheses that will be answered by monitoring, the sequence of scientific decisions will determine the monitoring approach required. These decisions will determine the most suitable monitoring design, data collection methods, as well as the analyses and interpretation of data. These are necessary considerations because the program objectives will differ between functional and effectiveness monitoring, specifically with regards to how they measure the effectiveness of management measures (Table 1). *Functional monitoring* may be a science-based, scaled-down version of effectiveness monitoring that relies on surrogate metrics to assess whether management measures provide expected conditions suitable for fish to carry out their life processes. The program objectives of functional monitoring are to determine if management measures recommended in the letter of advice or conditions outlined in the *Fisheries Act* authorization (e.g., mitigation, restoration, or offsetting) are functioning as intended (Table 1). Specifically, with a *Fisheries Act* authorization, an applicant is required to offset for habitat losses through restoration/enhancement, habitat creation, biological/chemical manipulations, and/or complementary measures. Monitoring of these authorization conditions to provide assurance that mitigation, restoration, and offsetting objectives were met, with regards to habitat function if not fully assessed via quantifying productivity, will be an important component of a functional monitoring program. In addition, where physical habitat is enhanced, restored, or created, quantifying the amount of functioning habitat before and after construction will be an important success criteria (Harper and Quigley 2005). Functional monitoring could employ rapid (e.g., < 0.5 days for two people to conduct the field sampling and < 0.5 days for two people for data management, analysis, and reporting; travel time not included) assessment techniques, using indicators of fish productivity (e.g., physical habitat measures, biological indicators) measured over the short-term (e.g., 1-3 years, but ultimate duration will depend on whether management objectives are met), and should provide measures of change in habitat function (Table 1). Rapid assessment could also be multiple, short site visits to install and remove data loggers, for example. Timelines for functional monitoring should be project-specific and may vary in the duration and start time; although, when possible, monitoring should take place before construction begins. Functional monitoring is not designed to determine losses or gains in fish productivity due to management actions, which is the main objective of effectiveness monitoring.

Effectiveness monitoring is defined as a scientifically defensible monitoring program that directly assesses the key metric (or indicator) of interest, and which must include the assessment of productive capacity (or surrogate) of the habitat compensation (DFO 2012). The program objective of effectiveness monitoring is to demonstrate that the management measures achieved the ecological milestones as set out in the terms of the *Fisheries Act* authorization (Table 1); milestones will be productivity-based and will in many cases be that there has been ‘no net loss’ of the productive capacity of the fish habitat (DFO 2012). It will be a research-based monitoring program where potentially a multi-year Before-After-Control-Impact (BACI) design will be used to evaluate the effectiveness of management measures (Pearson et al. 2005, Smokorowski et al. 2015). Measurements are often quantitative and directly linked to fish productivity (e.g., recruits-per-spawner, fish density, fish growth, fecundity). Effectiveness monitoring was the focus of a previous CSAS report (DFO 2012), while this report will focus on functional monitoring.

Table 1. Comparison of compliance, functional, and effectiveness monitoring program attributes based on the framework of DFO (2012).

Attribute	Compliance Monitoring	Functional Monitoring	Effectiveness Monitoring
Program Monitoring Objectives	Determine if the proponent executed the work as described in the letter of advice or Fisheries Act authorization	Determine if management measures are functioning as intended (e.g., if there has been a net change in the quantity of habitat, and if there has been a change in habitat function)	Determine if the management measures achieved the ecological milestones as set out in the terms of the *Fisheries Act* authorization (e.g., there has been a change in fish productivity)
Response type	No response	Indicator linked to fish productivity (e.g., quantity of suitable substrate for spawning)	Direct measure of fish productivity (e.g., fish density, abundance, biomass, growth, maturity)
Recommended Design	No design, audit of construction activities	Project-specific based on the questions monitoring will inform	Multi-year BACI
Application	All projects	Impact is small and/or well understood (e.g., road crossings, culverts)	Impact is large and/or complex and or has high uncertainty (e.g., gravel extraction)

Although both are quantitative approaches, functional monitoring would be a scaled-down version of effectiveness monitoring. Components that can be scaled-down are: 1) the complexity of the indicator, 2) the effort per survey, possibly by employing rapid assessment techniques, and 3) temporal and spatial replication, by reducing the number of control sites, years of monitoring, and the number of samples within a site.

The application of functional monitoring will be determined by the questions being addressed by the monitoring, and therefore there may be differences in the recommended design and indicators used in each application. We present hypothetical examples and possible monitoring details to highlight different applications of functional monitoring below and in Table 2.

Functional monitoring of habitat mitigation

Example: The placement of a culvert to mitigate the building of roads across streams and maintain an upstream and downstream migration corridor for multiple species and life stages of fish (Table 2). At a project level, the impacts of building roads on stream connectivity are relatively small and well documented compared to construction of dams or weirs (Ogren and Huckins 2015), making it suitable for functional monitoring. The specific objectives of functional monitoring for this project would be to determine if the culvert that was constructed matched or exceeded the specifications of the design and to determine if the culvert provides upstream and downstream passage to all fish species. For this example, a monitoring design that evaluated the connectivity of the habitat before and after construction would provide useful comparisons. Ideally, this would include sites located upstream and downstream that are sampled before and after construction. The challenge with monitoring culverts only immediately after their construction is that they tend to degrade over time; therefore, a staggered approach to monitoring (e.g., year 1 for compliance plus function and then five years after construction for function) may be the most effective approach. In many cases, data collected before construction are not available. In these cases, the indicator values measured after construction could be compared to species-specific benchmarks. Indicators could represent the physical habitat parameters that are related to passage requirements of each fish species (e.g., water velocity, water depth, gradient). It would be important to collect information about changes in presence of species either by comparing their presence before and after construction in the upstream site, or in the upstream and downstream reaches after construction. Sampling effort could be reduced by targeting key species (e.g., species with poor swimming performance), and/or sampling for presence/absence using a relative catch-based statistic (e.g., relative species-specific CPUE).

Functional monitoring of habitat restoration

Example: The transplantation of macrophytes (e.g., seagrass beds) in an estuary that was impacted by excessive sedimentation from channel dredging to restore nursery habitat function for salmonids (Table 2). Transplanting macrophytes is a common technique for restoring impacted macrophyte habitat that generally has well-known outcomes (Neto et al. 2013). The specific objectives of functional monitoring for this project would be to determine if the areal extent of the macrophyte habitat matched or exceeded the habitat before construction, as well as to determine if the habitat is suitable as a nursery habitat for fish. For this project, multiple monitoring designs could be used, including assessing the macrophyte habitat before and after construction; comparing the restored site to nearby reference sites in a desirable state (e.g., pristine, least-impacted), if available; or comparing the restoration site before and after, and to nearby control sites. Indicators could include the area of habitat restored, percentage cover of macrophyte, and density of macrophyte shoots.

Functional monitoring of habitat offset

Example: The addition of off-channel spawning habitat for salmonids to compensate for a channel diversion that led to loss of the same type of habitat (Table 2). The creation of off-channel habitat has been used in many locations throughout British Columbia to compensate for lost or degraded habitat (Pearson et al. 2005). The specific objectives of functional monitoring for this project would be to determine if the newly created habitat matched or exceeded the specifications of the design, and to determine if the habitat is suitable for spawning salmonids. For this example, a monitoring design that compares the habitat characteristics at the impacted site before construction to the newly created habitat, and to the impacted site after construction would provide useful comparisons. Favourable comparisons would suggest that the new habitat would support the fish productivity of the lost habitat. The newly constructed habitat could also be compared to nearby control sites in a desirable state (e.g., pristine, least-impacted), if available. Species-specific benchmarks of key spawning habitat parameters could be used to determine if the habitat is functional. Ideally, indicators would represent the physical habitat parameters that are related to spawning requirements of salmonids (e.g., substrate composition, cover, water velocity, water depth, oxygen, and temperature).

Table 2. Examples of when and how functional monitoring could be used to evaluate the success of mitigation, restoration, and offsetting measures.

Management Measure	**Habitat Function**	**Project Monitoring Objectives**	**Monitoring Design**	**Indicators Measured**
Mitigation – installation of a culvert to mitigate for impacts on fish passage due to road building.	Passage for multiple species of juvenile and adult fish	Compliance - Determine if the constructed specifications met or exceeded the design specifications; Function – does the culvert provide upstream and downstream passage?	Upstream and downstream assessments before and after construction (e.g., Before-After-Control-Impact design) where the assessment after construction is staggered (e.g., 1 and 5 years); No before data – use species-specific benchmarks to determine if physical habitat was limiting movement.	Presence of poor swimmers upstream, physical habitat characteristics (e.g., gradient, water velocity (m/s), water depth (m)).
Restoration – planting of macrophytes to restore habitat impacted from dredging in an estuary	Nursery habitat for fishes	Compliance - Determine if the required area and density of macrophytes was restored; Function – does the restored macrophyte area provide cover and food for fish?	Before and after construction assessments (e.g., Before-After design); Compare restored site characteristics to another site in a desirable state (e.g., Reference Condition Approach design); Compare restored site characteristics to the control site, both before and after construction (e.g., Before-After-Control-Impact design)	Area (m^2) of restored habitat, shoot density of macrophytes (shoots/m^2), % macrophyte area.
Offsetting – creation of off-channel stream habitat	Spawning habitat for salmonids	Compliance - Determine if the required area was constructed; Function – does the habitat provide cover and appropriate spawning substrate for adult salmon?	Before and after construction assessments (e.g., Before-After design); Compare offset site characteristics to another site in a desirable state (e.g., Reference Condition Approach design); Species-specific benchmarks used to	Area (m^2) of offset habitat, substrate composition, % undercut banks, % pools, oxygen, temperature, water velocity (m/s), water depth (m).

Management Measure	Habitat Function	Project Monitoring Objectives	Monitoring Design	Indicators Measured
			determine if physical habitat was within the requirements for spawning.	

Objectives

The objectives of this document are to present a review of monitoring designs, methods, and the types of information required for the development of a functional monitoring program. This review aims to address the following objectives:

1. Determine the recommended monitoring designs and methods for functional monitoring approaches based on monitoring objectives (e.g., rapid assessment techniques/select surrogates or indicators) to assess mitigation, offsetting, and restoration measures that are designed to reduce impacts to fish and fish habitat.
2. Identify information and analyses needed to support a science-based functional assessment of mitigation, offsetting, and restoration measures.
3. Determine the feasibility of gathering functional monitoring data using a checklist-style approach that can be applied consistently among project types and stages of construction (i.e., for each project type, a checklist of specific information to collect, with proponent-led monitoring and/or site visit during construction monitoring, and post-construction monitoring that can be applied in a consistent manner by FPP biologists).
4. If a checklist-style approach is considered feasible for various project types and stages of construction, identify the recommended fields.

To achieve these objectives, we provide a broad discussion about important considerations for designing a functional monitoring program, which is followed by a series of recommendations. We first present the concept of monitoring effort and the value of information. Introducing this concept provides a foundation for determining a reasonable amount of effort (effort per survey) that should be invested into a functional monitoring program. This is important to determine prior to the review because it provides context for our recommendations with regards to effort. While the level of effort we assume (presented in the sections below) may change between now and the implementation of this program, an introduction to these concepts will help guide the FPP in determining the appropriate levels. Next, we conducted a literature review to determine the state of functional monitoring programs employed around the world. We use the information gleaned from the review to summarize and discuss the monitoring designs, rapid assessments, analyses, and methods found for measuring habitat function. Finally, we discuss important considerations for developing a standardized, checklist-style functional monitoring program for aquatic habitats in Canada.

MONITORING EFFORT AND THE VALUE OF INFORMATION

Researchers are often faced with the difficult decision of how much effort to invest in a monitoring program. This is a primary consideration of any monitoring program and should be driven by monitoring objectives. The cost of acquiring the information can also be used to inform effort once the monitoring objectives have been met. We define effort as the amount of time and money to design, survey, analyze, and report on a science-based monitoring program. The level of monitoring effort is determined by the design of the program, including the type of comparator (e.g., Reference Condition Approach or BACI design), how many sites, how many times each site is visited within and among years, the number of indicators collected, the level of sample processing, type of analyses, repository of data, and type and extent of reporting. Beyond cost, for a well-designed program, effort also directly influences the reliability of information gained from monitoring through the inference drawn from the data collected and the certainty of management decisions based on that inference.

Figure 1 highlights the trade-off between the cost and benefits of acquiring information, which defines the value of information for a given effort. Costs are defined as the financial costs of acquiring information through monitoring, and the benefit is the certainty with which management decisions are made. If managers can identify reference points for the minimum and maximum amount of certainty that is required to make management decisions (effort driven by monitoring/management objectives), they can then be used to identify the minimum and maximum amount of monitoring effort to achieve management objectives. The value of information is defined as the difference between the cost and benefit for a given level of monitoring effort. The value of information can be used to identify reference points for the minimum, optimal, and maximum monitoring effort when costs are considered (effort driven by costs). The shape of these relationships will vary among monitoring designs; thus, the reference points can be used to compare the value of information derived from different monitoring scenarios. We expand on these concepts below.

This framework can be used to identify the level of monitoring effort required to meet a range of objectives: 1) the minimum amount of monitoring effort necessary for the benefit of monitoring to be justified by the cost; 2) the amount of monitoring effort that optimizes the cost-benefit ratio (i.e. the most efficient amount of effort); 3) the maximum amount of effort that can be justified by the cost; 4) the minimum amount of effort required to make a management decision; and 5) the maximum amount of monitoring effort, where no further information will alter the management decision.

Minimum effort and diminishing returns on the value of information

There are lower and upper bounds to the level of monitoring effort required for making good management decisions. For example, there is a minimum level of information that is required for the information to be useful in achieving the monitoring objectives. This could be thought of as the minimum monitoring effort. Visiting a field site and taking one measurement is unlikely to yield useful information. At the other end of the spectrum, there is a theoretical maximum amount of effort that should be invested into a monitoring program, and after that point minimal further value (i.e., information) can be gained. In the context of decision making, the minimum effective effort could be thought of as the amount of information required to influence the decision of a manager, while the maximum effective effort is the amount of information at which further data collection would not change the decision or reduce uncertainty. If a cost-effective program was an objective, the value of information (VOI = benefit – cost) could also be used in identifying the minimum, optimum, and maximum monitoring effort. For example, there is a minimum level of effort required for the value of the information collected to justify the cost (e.g., VOI > 0). The optimum identifies the most cost-effective level of effort. At the other end of the spectrum, there is a theoretical maximum amount of effort that should be invested into a monitoring program, and after that point no further value (i.e., information) can be justified against the acquisition cost (e.g., VOI ≤ 0). The minimum and maximum cost-effective effort should only be considered if that level of effort is within the range of effort required for making good management decisions. Taken together, these points suggest there is a window of effort that is effective for making decisions and within that window, an optimal effort that is the most cost-effective (Figure 1A). While these are theoretical bounds on effort as it relates to meeting monitoring objectives, they serve a practical purpose to help guide program biologists when considering how much effort to invest in monitoring.

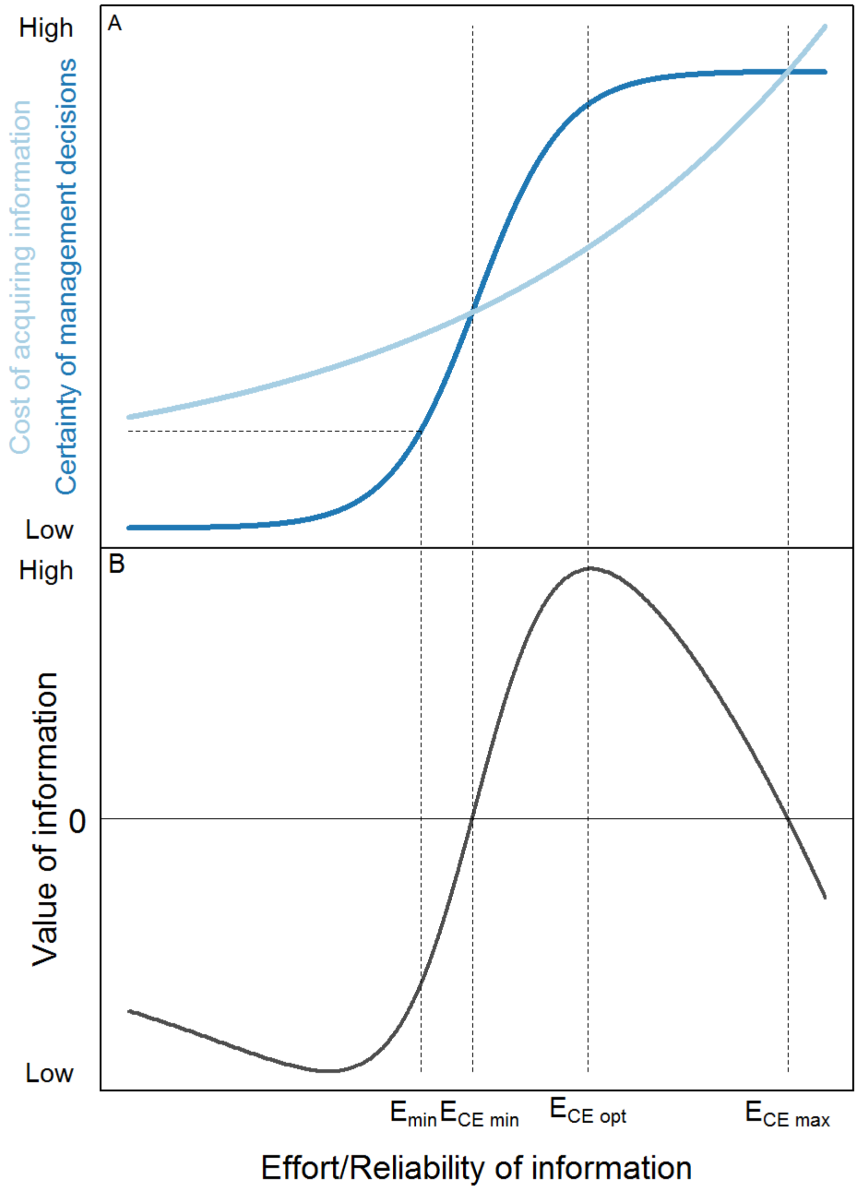


Figure 2. A) The relationships between effort and the cost of acquiring information (light blue line), and the effectiveness of using the information (certainty of management decisions) (dark blue line). B) The relationship between effort and the value of information. The difference between the cost and effectiveness represents the value of information, which is the $E_{CE\ min}$, $E_{CE\ opt}$, $E_{CE\ max}$ and represents the minimum, optimal, and maximum cost-effective monitoring effort, respectively. E_{min} is the minimum effort required to make a management decision. When monitoring effort is too low (Effort < E_{min} & $E_{CE\ min}$), there is no value of the information gathered. As effort increases past the optimal point, eventually there are diminishing returns on the value of information (Effort > $E_{CE\ opt}$) to the point where the gain in effectiveness is no longer justified by the costs. Between these bounds is a window of cost-effective monitoring effort. Theoretical representations of the effort-information trade-off such as this need to be placed in the appropriate context of how inference will be drawn about impacts on habitat or how the information will be used to make decisions.

Figure 2 assumes that the amount of information collected and the value of that information for making management decisions are proportional. For this assumption to be valid, the monitoring objectives must be clearly defined, matched by data collection at the appropriate scales, and relevant to the management decisions that will ultimately be made based on the inference drawn from the data. Matching data collection to monitoring and management objectives is crucial for the success of any monitoring program and should be paramount during program development.

LITERATURE REVIEW

We conducted a literature review of aquatic functional assessment to collect information that can be applied to the development of a functional monitoring program in Canada. We focused on the global use of functional monitoring approaches, assessment methods used, and the effectiveness of those monitoring programs. The following factors were considered in our review.

Monitoring design: Functional assessments are applied using a range of monitoring designs (Figure 3). Most existing monitoring approaches use one of the five designs shown in Figure 3. These are:

1. Before-After-Control-Impact (BACI) - Best practice for determining the function or effectiveness of an intervention (mitigation, restoration, and offsetting) is a BACI design (Green 1979, Stewart-Oaten et al. 1986, Underwood 1991, 1994, Stewart-Oaten and Bence 2001). At minimum, this design compares an altered site to a control site before and after the habitat has been altered (e.g., restoration or impact) (Figure 3A, B). Typically, data are collected for multiple control sites and multiple years before and after the alteration. While this study design is considered the most scientifically defensible, the effort associated with this level of data collection may be impractical for many functional monitoring projects.
2. Control-Impact (CI) – The CI design is similar to the BACI design, but only uses spatial controls to compare to an altered site (Figure 3C, D). Data are collected after the change in habitat. The most rigorous application of this design would include multiple control sites. Control sites can be reaches or areas within the same system (e.g., stream, estuary shoreline), or can be from nearby systems that represent similar conditions.
3. Before-After (BA) – The BA design is similar to the full BACI design, but only uses temporal controls to compare to an altered site (Figure 3E, F) (Underwood 1991). Data are collected before and after the change in habitat. The most rigorous application of this design would include multiple years of sampling before and after the habitat change.
4. Reference Condition Approach (RCA) – The RCA compares an altered site to a set of conditions defined by multiple reference sites that represent some desirable state (e.g., undisturbed, pristine, or not-impaired) (Figure 3G, H) (Stoddard et al. 2006). Reference sites are often selected to match hydrogeomorphological characteristics of the altered site that are often resilient to change by human activities. This reduces the potential for natural variation to be the driver of any observed differences (Barbour 1998, Perrin et al. 2007). Reference conditions and comparisons between reference and altered sites are made using multivariate analyses. Data are collected after the change in habitat at the test site; reference site data can be collected before or after changes to the altered site habitat. The most rigorous application of this design would include multiple reference sites. Often the number of reference sites is much greater than the number of controls present in a BACI, CI, or BA design.

5. Normal Range Approach (NRA) – The NRA is similar to the reference condition approach, but uses conventional statistical procedures to compare test sites to the normal ranges of data (i.e., reference conditions). The status of a test site can be based on whether the site values are outside the normal range or based on pre-defined benchmarks derived from various sources (e.g., models, experiments, data distributions). The normal range approach is flexible in that it can incorporate the spatial and temporal aspects of other monitoring designs (e.g., BA, CI, BACI, or RCA).

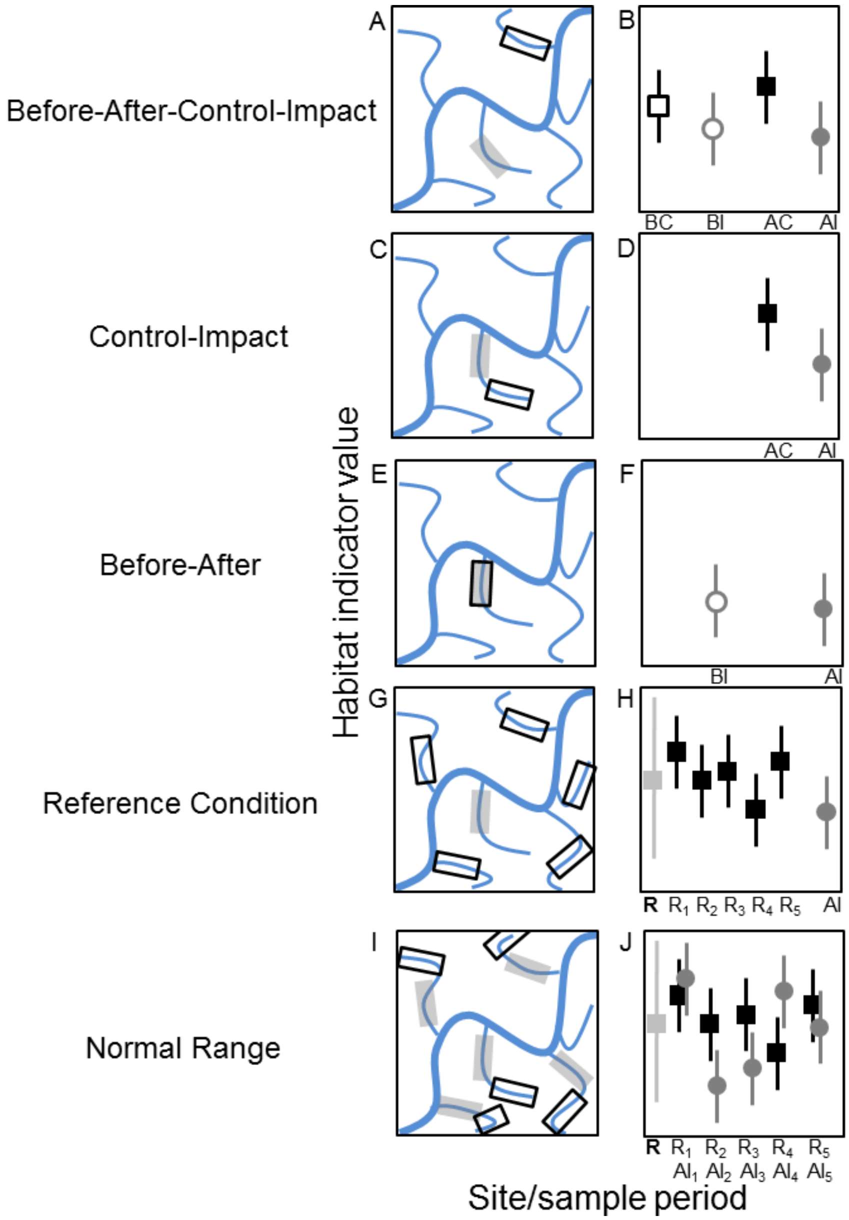


Figure 3. First column (A, C, E, G, I) shows the spatial organization of sites for five commonly used monitoring designs in a network of streams. Transparent grey areas represent altered sites and the black boxes represent the control (A, C, E) or reference sites (G, I). The second column (B, D, F, H, J) shows hypothetical data and how comparisons are made between control and altered sites, and/or sampling periods. Grey points represent altered sites/samples and black points are the control sites/samples, open points represent samples taken before the habitat alteration and closed points represent samples taken after the habitat alteration. The light grey squares represent the distribution of values observed in reference sites.

The level of spatial and temporal replication varies among different monitoring designs. Best practices would consider *a priori* the spatial and temporal variability in the data being collected and the potential effect size of the impact when designing the study. For example, measurements that are temporally stable are favoured over others that vary seasonally, especially when timing of data collection cannot be controlled.

Rapid Assessments: Rapid data collection is a key component of functional monitoring. Assessments that can be conducted quickly will allow for more sites to be monitored. There are several definitions of "rapid" assessments, complicating the use of the term. Definitions vary in the amount of time and the assessment components that are included (e.g., planning, data collection, laboratory analysis, data management, reporting). Many of the rapid assessment protocols use biotic indicators where only the field data collection component of the assessment is rapid.

Indicators: Functional monitoring will measure habitat indicators (indirect measures) of fish productivity to determine if the habitat is functioning properly. These measures are expressed through a range of metrics that are designed to capture different habitat functions. Measures for functional monitoring need to be: 1) sensitive to environmental change, 2) exhibit low natural temporal and spatial variability, and 3) quick to measure, process in the lab, and/or analyze (i.e., rapid) (Rice 2003). There is no one metric or suite of metrics that fit all monitoring scenarios, thus variability in metric performance should be tested during the development or early stages of the monitoring program to ensure the protocols will provide sufficient information for detecting a change when change has occurred, or no change when it has not.

Standardized monitoring: Agencies throughout the world use standardized monitoring programs to evaluate the function of the aquatic ecosystems (Appendix A). For example, standardized protocols for collecting and analyzing benthic macroinvertebrates in streams and lakes are used in Canada, USA, Australia, and the UK. These are large-scale programs applied regionally, often designed to measure ecosystem health. The USEPA has developed a rapid bioassessment protocol, quantifying metrics such as physical habitat (e.g., substrate, riparian vegetation disturbance, and cover for fish), macroinvertebrate and fish assemblages (e.g., diversity metrics), to assess the health of wadeable streams throughout the USA (Barbour 1998). While we did not find any global assessments of performance of large-scale monitoring programs, there are regional assessments that provide valuable insight into how well they perform. Existing standardized protocols have been adopted by several other countries, and there continues to be development of these monitoring programs, further advancing the field of standardized aquatic monitoring (e.g., Bailey et al. 2014). Some benefits of standardization of monitoring programs are protocols to ensure metrics are measured correctly, consistency in indicators measured and data entry, determining the appropriate analyses prior to data collection, and clear reporting of results. They can also provide broader benefits to the science around mitigation, restoration, and offsetting activities through meta-analysis of program results that evaluate the effectiveness of different management actions. Meta-analysis is a powerful analytical approach that can be used to determine the overall effects of a given management measure by assessing the effects observed in multiple individual projects (Arnqvist and Wooster 1995).

The literature review examines functional monitoring approaches for both freshwater and marine habitats. The review focuses on higher-level monitoring concepts (e.g., monitoring design, metrics for monitoring function, and considerations for standardized monitoring) that are transferable between ecosystems and among system types. Examples of potential indicators by habitat type are provided, but the review does not provide recommendations for specific indicators or metrics that should be measured. Indicators are likely to be ecosystem- and system-specific. We discuss metric characteristics and what might constitute a better class of metrics for functional monitoring, and provide some examples. Important differences between marine and freshwater ecosystems (e.g., estuary and marine coastal vs. streams, rivers, and lakes) with regards to functional monitoring and metrics selected will be highlighted, where they exist.

REVIEW METHODS

We conducted a literature review to determine the state of knowledge on functional assessments for evaluating aquatic habitat function. We followed key components and principals of a systematic review (Evidence Collaboration for Environmental 2018), but limited documentation steps to reduce the amount of time required to complete the review. We followed four steps that included: 1) developing a list of search terms and Boolean operators, 2) searching literature databases using the final search term list, 3) inclusion criteria for determining if studies would be included for further evaluation, and 4) extracting data. We extracted data from studies to summarize where functional assessments are conducted, what monitoring designs are used, the types of data collected, if they use a standardized protocol, the level of replication, and whether or not they were rapid. Although we did not include reviews, methods, or protocol papers in our formal assessment, relevant papers within these categories were reviewed and considered during the writing of the report.

Search Terms

The base list of search terms was developed through consultation with the steering group and review of systematic protocols on similar topics. The final list of search terms was developed through an iterative process where the inclusion and exclusion of terms and Boolean operators were tested (Table 3). Results from each set of search terms were examined by identifying if the search results included key papers identified *a priori* or during the searches, and the number of results generated. If changes to the search term list resulted in the loss of key papers, the list would revert back to the previous version. Similarly, if results for a set of search terms was too high (i.e., > 10 000), we reverted back to the previous list.

The final list of search terms was placed in four broad categories: 1) study objective – what is the key focus of the paper (e.g., identifying indicators, or examining fish or invertebrate habitat); 2) habitat/population – habitat needed to be aquatic; 3) intervention – what type of intervention was involved; and 4) Assessment – how was the mitigation activity or habitat assessed.

Table 3. Terms and Boolean operators used to search Web of Science and Scopus literature databases.

Study objective	(Fish* OR Invertebrate* OR Index OR Indices OR Indicator* OR Surrogate OR Proxy)
	AND
Habitat/population	(Habitat* OR Marine OR "Fresh water" OR Freshwater OR Aquatic * OR Stream* OR River* OR Wetland* OR Marsh* OR Lake* OR Estuary* OR Reef* Or “Near Shore” OR “In Shore” OR Coast*)
	AND
Intervention	(Mitigat* OR Offset* OR Restor* OR "No-Net-Loss" OR Stressor*)
	AND

Assessment	(Checklist OR “Habitat assessment” OR “Rapid Assessment” OR "Visual Assessment" OR “Functional Monitoring” OR “Offset Monitoring” OR “Restoration Monitoring” OR “Effectiveness Monitoring” OR Effectiveness OR "Biological Function*" OR "Ecological Function*" OR "Before-After" OR BA OR "Before-After-Control-Impact" OR BACI OR "Before-After-Impact-Control-Paired" OR BACIP OR "After-Control-Impact" OR ACI OR "Reference Condition*")

Search terms were used to search several literature databases. Our main search focused on Web of Science and Scopus (Figure 4), as these are the primary science-based databases of peer-reviewed literature. Secondary searches focused on provincial and federal government sites. We searched nine federal and provincial databases using eight common search terms, as these websites did not allow the use of a large multi-faceted search string.

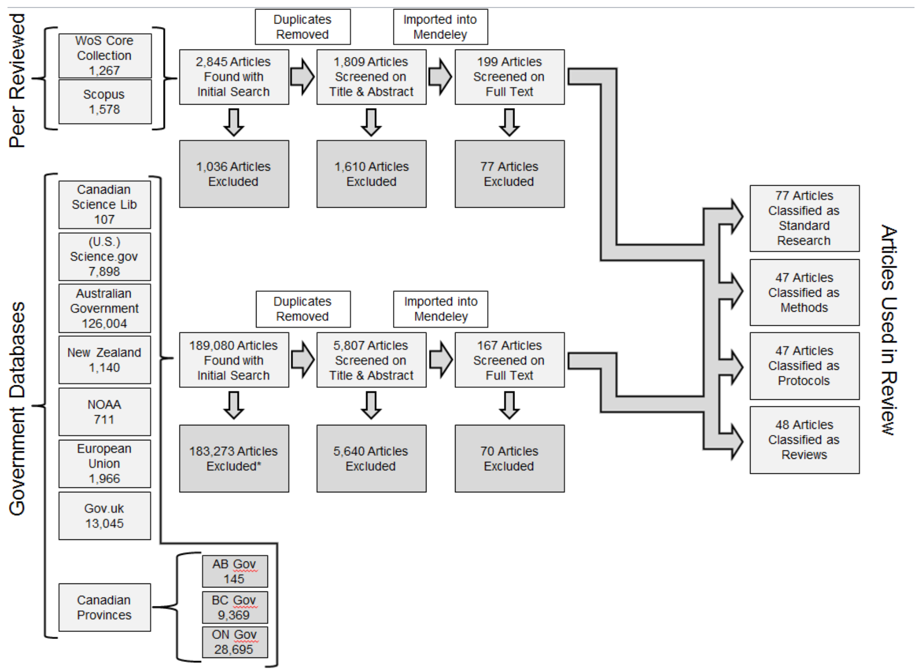


Figure 4. Literature review work flow, number of search results, and screening results from peer-reviewed literature (Web of Science and Scopus) and government websites and databases. All peer-reviewed results were screened, however, due to the large number of results and low relevance of most papers found on government websites and databases, we limited the number of results reviewed through setting page and number limitations. The reduced search term list was used because most websites did not allow multi-faceted search strings. The reduced search term list included: Aquatic Rapid Assessment Method, Marine “Rapid Assessment” Monitoring, Rapid Bioassessment Protocol, Near Shore Monitoring, Marine “Habitat Assessment”, Rapid Assessment Protocol, Estuary Monitoring, and Rapid Lake Assessment. Urls for all the websites and databases can be found in Appendix C.

Each paper was first screened at the title and abstract level for relevance to our topic (Figure 4). Papers that met the inclusion criteria (Appendix B) were then downloaded and imported into the reference manager software Mendeley for review at the full text level. Briefly, papers were included if they were about aquatic habitat used by fish for spawning, rearing, foraging, or migration. Papers also must have evaluated a human-related stressor, or a habitat-related management action (e.g., mitigation, restoration, or offsetting). Furthermore, they needed to have used rapid assessment techniques or short-term monitoring. This criterion was used to avoid overlap with long-term effectiveness monitoring. All papers that met the above criteria were categorized as standard papers and included in the analytical component of the review. Studies that did not meet the criteria but provided useful information on the topic were categorized as reviews, methods, or as protocols. The inclusion criteria were applied at the title and abstract level, but if the information needed to assess the paper at that level was not available, papers were put into a category for further review at the full text level. All papers that were screened in or flagged for further review during the title and abstract review were then screened at the full text level using the same criteria. The reviewers used their discretion when reviewing a paper that may not have fit all the inclusion criteria but would make a valuable contribution to the review. We also included papers that were identified during targeted searches or exploration based on references or papers sent to us from colleagues.

DATA EXTRACTION AND SUMMARY

To capture how and where functional assessments were used, we extracted and summarized data from the standard papers. This included numerous fields about where the research was conducted, the type of ecosystem and habitat, the research topic, the monitoring design used, the level of replication, and whether or not the assessment methods were rapid. Rapid assessment was defined as an assessment that requires < 1 day to complete the monitoring (e.g., < 0.5 days for two people to conduct the field sampling and < 0.5 for two people for data management, analysis, and reporting); travel time is not considered in this definition. The data for standard papers were collected in an Excel spreadsheet (Appendix A) along with the meta-data for each field. We extracted and reviewed 219 papers at the full text level, and placed each paper into one of four classifications: 1) standard research, 2) methods, 3) protocols, and 4) reviews. Only the standard research papers met our inclusion criteria and were included in our summary results. We present simple summaries of the data by highlighting important components of the dataset and commonly used methods or approaches. These results are presented in the discussion of each of the main topics. In addition to simple summary statistics, the literature review provided most of the discussion points and references presented in the following sections.

MONITORING DESIGNS

From our review, we found there is a range of monitoring designs used in environmental impact assessments and assessments of the effectiveness of restoration or offsetting habitat. Approaches can be grouped into four main categories, as previously described: 1) Before-After-Control-Impact, 2) Before-After, 3) Control-Impact), 4) Reference Condition Approach, and 5) Normal Range Approach (Figure 2 and Table 4). These categories represented 72% of the studies reviewed. While there are other monitoring designs that could be used, such as trend-by-time analysis (Wiens Parker 1995) and iterative normal ranges (Arciszewski et al. 2017), they often require long time series of data and are therefore less relevant to functional monitoring, which will focus on shorter time periods of data collection that may not be continuous (DFO 2012).

All monitoring designs have benefits and limitations (Munkittrick 2009). In this section, we summarize the most common monitoring designs found in our review. The design is introduced, benefits and challenges are outlined, and the types of analyses are discussed, along with the level of effort required for implementation. Notably, the majority of examples focus on freshwater, due to the lack of marine-focused standardized assessment protocols that came up during our literature searches.

Table 4. Monitoring designs commonly used in the assessments of habitat alteration (e.g., impacts, habitat mitigation, restoration, and offsetting.)

Design	Temporal Replication	Spatial Scope	Benefits	Challenges	Effort	Most Applicable	Least Applicable	Key Considerations
BACI	Often multiple years of data collected before (temporal control) and after habitat alteration	Often multiple control sites assessed - recommend 3-5 sites	Controls for variation in both the control and altered site	Requires adequate temporal replication	Highest effort design at the time of assessment; no further costs once assessment is complete	Most applicable when appropriate spatial and temporal replicates are available	Least applicable when alteration and response are tightly linked and/or response's temporal distribution is uniform	Must sample impact site prior to impact – preplanning necessary; appropriate temporal, spatial replication, and control sites
Before-After	Often multiple years of data collected before (temporal control) and after habitat alteration	No spatial controls	Can be used to assess a site where appropriate control sites are limited	Requires pre-impact data; Length of time for pre/post data collection Cannot control for temporal variation not attributable to the impact	Less effort than a full BACI design; no further costs once assessment is complete	Most applicable for projects performed in unique environments (e.g., very large rivers, large inland deltas) for which controls sites are limited	Least applicable for projects that are under stringent timelines for which the before data collection is unfeasible	Must sample impact site prior to impact – preplanning necessary
Control-Impact	No data collected before the habitat alteration, data only	Often multiple control sites assessed -	Can be used to assess a site with no	Does not control for temporal variation not attributable	Less effort than a full BACI design;	Most applicable when there is an immediate monitoring need that does not allow for the	Least applicable when systems are highly variable or sites are unique	Appropriate control necessary – can be difficult for

Design	Temporal Replication	Spatial Scope	Benefits	Challenges	Effort	Most Applicable	Least Applicable	Key Considerations
	collected after the habitat alteration	recommend 3-5 sites	baseline data	to the impact	no further costs once assessment is complete	collection of baseline data, and high spatial replication is possible		more unique sites
Reference Condition Approach	Often no temporal replication, but temporal replication possible	Multiple control sites required, recommend 10 reference sites per reference group	Can be used to assess a site with no baseline data	Does not capture the baseline conditions of the impacted site; can be difficult to adequately represent reference condition	Lowest effort design at the time of assessment; additional costs of monitoring reference sites may be substantial	Most applicable for evaluating ecosystem health at large spatial scales	Least applicable for assessing local habitat changes (e.g., fish passage, availability of structure and cover)	Appropriate number and type of reference sites; costs may be alleviated by using existing information to develop indicator benchmarks
Normal Range Approach	No data collected before the habitat alteration, data only collected after the habitat alteration	Data may be used from multiple sites	No baseline data required	High data needs (e.g., long time series) to identify normal ranges or benchmarks. Normal ranges may be affected by shifting baselines.	High effort to establish normal ranges or benchmarks unless models or data exist in the literature; lower effort needed for each project.	Most applicable where there is a strong linkage between response and habitat alteration, and benchmarks and normal ranges can be defined using laboratory or field data, experiments, or models	Least applicable when there is little information about the system and relationships with the response	Transfer of information on normal ranges or benchmarks among projects/ systems/species

Design	Temporal Replication	Spatial Scope	Benefits	Challenges	Effort	Most Applicable	Least Applicable	Key Considerations
Trend-by-Time Analysis	No data collected before the habitat alteration, data only collected after the habitat alteration, multiple years required	Data may be used from multiple sites	No baseline data required	Long time series required, assumes dynamic equilibrium	High effort to establish time series.	Most applicable when there is an immediate monitoring need that does not allow for the collection of baseline data	Least applicable when systems are highly variable or where long-term data collection is not feasible	Evaluates relationship between metric and impact (continuous)

BACI, BA, CI DESIGNS

A well designed Before-After-Control-Impact study is considered the gold standard for assessing environmental impacts or habitat alterations (Stewart-Oaten et al. 1986, Underwood 1991, 1994, Smokorowski and Randall 2017). However, we found that the BACI monitoring design is typically applied to long-term research projects that would be used in effectiveness monitoring. For example, only 13% (10/77) of the studies reviewed evaluated a BACI design. Before-After and Control-Impact designs were also not well represented in the literature, consisting of only 5% (4/77) and 9% (7/77) of the studies we reviewed, respectively.

The strengths of the BACI design are that it controls for spatial (Control-Impact) and temporal (Before-After) variation, and their interaction. For a site to be deemed altered there must be a change in the habitat from before to after the construction, and that change must be different than what is observed in the control site. It is this interaction that is central to the strength of the BACI design (Figure 5).

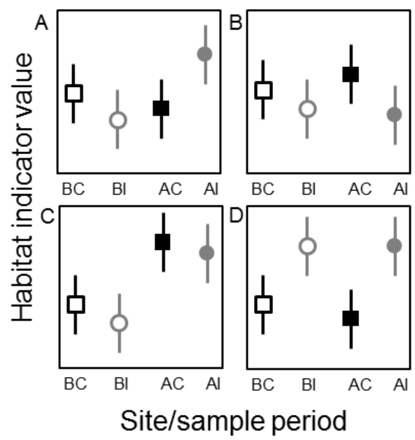


Figure 5. Hypothetical results from a BACI design applied to a restoration project, where squares represent control sites, circles represent altered sites, open symbols represent before, and closed symbols represent after construction. A) Shows a significant interaction between space and time, where a difference is observed in a habitat indicator between before and after construction in the altered site but no difference is observed between before and after at the control site. This result suggests that the change in habitat was related to the restoration. The same conclusions would have been made had the design only considered the before and after or the control and impact. B) Shows no significant differences among space or time. The same conclusions would have been made had the design only considered the before and after or the control and impact (i.e., restored site). C) Shows a significant effect of time but not space. This means that differences in the habitat indicator values were observed before and after construction at both sites suggesting that the change in habitat was not related to the restoration. The same conclusion would have been reached if a control-impact design was used, but a different conclusion (an effect of restoration) would have been reached if a before-after design was used. D) Shows a significant effect of space but not time. This means that differences in the habitat indicator were observed between the restored and control sites both before and after construction, suggesting that there was no change in habitat and that differences between sites existed before construction. The same conclusion would have been reached if a before-impact design was used, but a different conclusion, which would have supported an effect of the restoration, would have been reached if a control-impact design was used.

The BACI design is an intensive study design. For instance, we found that, on average, the ratio of the number of control sites to impact sites was 1.5 for BACI and CI designs, compared to 0.6 for studies that used the reference condition approach. While the level of rigor associated with a BACI design is appropriate for the assessment of projects that are expected to have a large impact (Smokorowski et al. 2015), it may not be the most cost-effective approach for smaller-scale projects expected to have a low impact, or when the impact is well understood (e.g., culvert placement, or small-scale restoration projects). It was estimated that the lab costs for enumeration of benthic macroinvertebrates and associated habitat measurements in a small impacted stream in British Columbia using the reference condition approach (see below) would cost $625, while the additional samples required when using a BACI design would cost more than $4000 (Perrin et al. 2007). However, this comparison ignores the additional costs of developing and maintaining the reference sites used to derive the reference condition for the RCA.

Before-After and Control-Impact designs are variants of the BACI and are often considered inferior because they fail to account for temporal (Before-After) and spatial (Control-Impact) variability that might lead to observing a change when it is not present (Type I error) or not observing a change when it is present (low power) (Figure 5). However, this assumes that a BACI is well designed, with control systems that exhibit correlated dynamics with the altered site; these assumptions are often not met. When controls are not correlated with the altered site, a Before-After design may be more powerful because the uncorrelated dynamics can lead to unexplained variability (Bradford et al. 2005). These study designs may be more appropriate for functional monitoring because they are logistically simpler to conduct and may be the only option if the appropriate controls do not exist (Before-After) or if pre-construction data are not available (Control-Impact).

Analyses for estimating impacts on habitat using BACI, BA, and CI designs are relatively simple. The main approach is to conduct a two-way analysis of variance (ANOVA) that includes the main effects of time (before and after the impact) and space (impact vs. the control site), as well as an interaction between time and space. The main effect of time indicates a difference between before and after (Figure 5B), and the main effect of space indicates a difference between the control and altered site (Figure 5C). However, the key parameter is the interaction between site and year, which indicates the altered site changed over time, while the control site did not (Figure 5A). For BA or CI designs the key parameters are simply the main effects of time or space, respectively. Significance of parameter estimates are often determined using significance (i.e., alpha) values of < 0.05; however, although this is ubiquitous in classic hypothesis testing, it is an arbitrary value that may not be appropriate for all monitoring programs (Mudge et al. 2012, Bradford et al. 2017). For unbalanced designs, alternative approaches (e.g., Restricted Maximum Likelihood) can be used to account for the differences in the number of years before and after the impact (Robinson 1987).

An alternative approach is to calculate the difference in the values between the control and altered sites and then conduct a two-factor test (between before and after). A significant parameter will indicate that the control and altered sites changed differently from before the impact to after. Smokorowski and Randall (2017) compared the results of their data using the two-way ANOVA, and difference between control and impact using the t-test. The t-tests indicated differences in invertebrate communities before and after the changes in flow, but no differences were indicated by the two-way ANOVA. Regardless of the statistical approach, it is important to note that classic hypothesis testing may not be the most appropriate way to statistically evaluate change in habitat function (Munkittrick 2009, Mudge et al. 2012) (see Considerations For All Designs section).

REFERENCE CONDITIONS AND BENCHMARKS

Reference conditions and benchmarks provide values that represent an expected condition (often from sites where significant human disturbance is absent or that are minimally disturbed) that can be used to make assessments of the ecological status of an altered site (Stoddard et al. 2006). These values are often derived from a distribution of values measured at reference sites but may also be derived from experiments or models. General approaches that apply the concept of reference conditions and benchmarks are the Reference Condition Approach (RCA) and 2) the Normal Range Approach (NRA). Biological and physical data can be used with either approach; however, the Reference Condition Approach is generally used with biological assemblage data, in particular benthic macroinvertebrates, whereas the Normal Range Approach is used with simpler biological and physical indicators. The main difference between the two approaches is in the analytical framework used to determine the reference condition or benchmarks. While there are other approaches that use the concept of reference conditions (Hawkins et al. 2010), we highlight the most commonly used approaches.

Reference Condition Approach

The Reference Condition Approach is one of the most widely used monitoring designs for environmental monitoring and was the most common monitoring design, used in 48% (33/77) of the studies reviewed. This monitoring design is applied at the national level for many countries and unions, including Canada, the United States, Australia, New Zealand, and the European Union (Appendix A), and can be used with both physical habitat and biological assemblage data. This monitoring design compares sites that have undergone habitat alterations (e.g., positive change – restoration; negative change – impacts from industry) to reference conditions. Reference conditions can be developed by measuring reference sites in pristine, or the least-impacted, condition.

Reference sites are typically monitored within a region and represent a broad range of site conditions from which categories related to impairment or function (Figure 6) are determined. High spatial replication of reference sites is the foundation of an effective Reference Condition Approach. Another approach uses long-term ongoing monitoring of protected areas as reference sites. An example of this is the National Estuarine Research Reserve System, which is a network of 28 protected areas that represent a range of biogeographical regions throughout the United States (Imperial and Hennessey 1996). This regionally-based approach could be useful for when larger systems are being evaluated and few appropriate reference sites exist.

Comparisons between test (altered) and reference sites requires that altered sites match the reference conditions predicted by reference sites with similar hydrogeomorphological characteristics (Barbour 1998, Perrin et al. 2007). These characteristics are selected because they are not influenced by human activities (e.g., gradient, catchment size, stream order). This process of organizing reference sites into reference groups provides a natural range of conditions and can identify suitable comparisons between reference and altered sites, eliminating the need to find the perfect control site. However, distinct reference groups are required so that classification errors related to the assignment of altered sites to reference groups do not reduce the power of the test between altered and reference site status.

The RCA provides operational flexibility in that the reference sites can be added and maintained independently of altered sites where the timing of monitoring is determined by *Fisheries Act* authorizations. If a regional set of reference sites exists, only the altered site needs to be measured by DFO or the proponent. This would allow agencies to easily and quickly react to monitoring needs associated with authorizations because there would be no planning for reference sites during the authorization process.

There are two main analytical components of the RCA analyses when applied to biological assemblage data (periphyton, macroinvertebrates, and fish). Briefly, a classification analysis is used to group reference sites based on their biota. There are important trade-offs for determining the organization of reference sites into groups. Organizing reference sites into many smaller reference groups with more restricted ecological attributes will increase sensitivity to disturbance (decrease Type I error), but can increase classification errors (assigning a test site to the appropriate reference group); whereas fewer larger reference groups with broader ecological attributes will potentially decrease sensitivity to disturbance (increase Type I error), but will lower classification error rate (Perrin et al. 2007). A discriminant function analysis is then used to build a model from the reference data that assigns the altered site to a reference group, based on the impact site's environmental variables. Once an altered site is assigned to a group, a comparison of the impact site's biota is made to the reference community (Figure 6). This analytical approach (or some variant) is used in the UK's RIVPACS (River Invertebrate Prediction and Classification System), Australia's AUSRIVAS (Application of the Australian river bioassessment system), and Canada's BEAST (Benthic Assessment of Sediment). Methods generally differ in whether they use presence-based taxonomic metrics (RIVPACS, AUSRIVAS) or abundance-based metrics (BEAST). Uncertainty in assigning sites to impairment level categories can be represented by probability ellipses in ordination space (Perrin et al. 2007). The ellipses surrounding the reference site coordinate values roughly represent their distribution. For example, if an altered site's ordination coordinates place it within the 90% ellipse, where 90% of the reference site coordinates are within this region, it would be categorized as not stressed. If a test site's coordinates lay outside of this region, where very few of the coordinates for reference sites exist, it would be categorized as impaired.

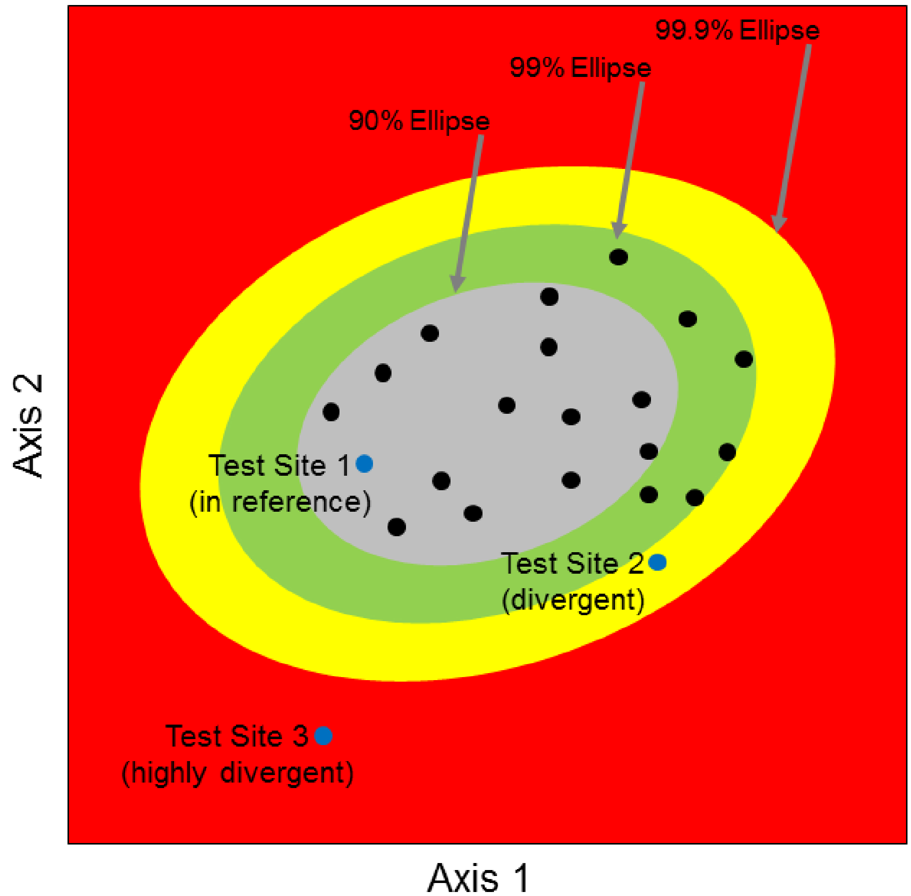


Figure 6. Hypothetical distribution of 19 reference sites (black circles) and three test sites (blue circles) in multivariate ordination space. Sites located near the centroid (center of axis 1 and 2) are the most similar to all other sites. Coloured ellipses are generated from reference site data and represent the 90th, 99th, and 99.9th percentile of the distributions, each corresponding to different levels of similarity with respect to the centroid. For example, Test Site 1 is located in the grey region of the plot, indicating that it is within a region of ordination space that includes 90% of the reference sites. This means that it is within the reference distribution and is considered "in reference condition". Test Site 2 is located in the yellow region, indicating that it is within a region that includes 99.9% of the reference sites and is considered "divergent" from the reference condition. Finally, Test Site 3 is located in the red region, where < 0.01% of reference sites would be found, and is "highly divergent" from the reference condition.

In a rare test of the performance of the statistical approaches for the RCA, and commonly used in national programs, Bailey et al. (2014) found that conventional methods resulted in higher uncertainty (lower Type I and Type II errors) as to whether an altered site was in reference condition when compared to several newer approaches (e.g., machine-learning, Bayesian approaches). They tested these statistical approaches using three large regional datasets of macroinvertebrate assemblages from wadeable rivers in Australia and the Yukon, and nearshore lotic sites in the Great Lakes. Current methods used by the Canadian Aquatic Biomonitoring Network (CABIN) program (Benthic Assessment of Sediment) were applied to the Yukon Territory dataset, and resulted in high Type 1 error rates (determining a site is in reference condition when it is not) and low power for detecting the severity of the impairment. Specifically, 30% of sites that were determined to be in reference condition were not; and only 35%, 40%, and 45% of mildly, moderately, and severely impaired sites, respectively, were appropriately categorized as such. The machine learning methods developed by Feio et al.

(2014) resulted in a lower Type 1 error rate of 15% and increased power for assigning impairment, whereby 25%, 68%, and 78% of sites were correctly determined to be mildly, moderately, and severely impaired, respectively. Many of the analytical improvements focus on the matching of reference sites to altered sites. Bailey et al. (2014) recommend that large-scale bioassessment programs, such as CABIN, AUSRIVS, and RIVPACS adopt newer analytical methods. Furthermore, evaluating across three different datasets demonstrated the variability in performance of all methods, and highlights that caution should be used when applying these methods at national and regional scales.

Although the Reference Condition Approach is the most commonly used approach for large-scale standardized monitoring programs, we identified several commonly recognized challenges to using the RCA to assess the function of aquatic habitats:

1. Identifying appropriate reference sites. Brinson and Rheinhardt (1996) recognized this as a major challenge. Selecting appropriate reference sites is difficult when assessing large water bodies such as large rivers, lakes, or estuaries because there are few suitable comparators within a region. For example, there are only 11 rivers in North America that meet the definition of a great river (Angradi et al. 2009), all located in different regions. Alternative approaches to reference site selection may be required. This could include longitudinal sampling least-disturbed sites within the river's watershed. The network structure of rivers will lead to inherently correlated values among sites, and a spatial analysis may be needed to deal with the lack of independence among sites. Another alternative would be to use benchmarks developed from existing data sources. Highly impacted ecoregions will also be difficult to assess using the RCA because there will be few pristine or minimally impacted sites. For example, the Lake Erie-Lake Ontario ecoregion is the most diverse ecoregion in Canada, but 78% of the land has been developed (Crins et al. 2009). To deal with this issue, others have altered the definition of a reference site to "best available" or "least impaired". This has important implications for how impacted sites are assessed. Similar challenges of selecting reference sites may occur in the assessments of coastal marine areas associated with highly impacted areas (e.g., Vancouver, Halifax).
2. Cost of developing and updating reference sites. A set of reference sites that can be used to compare with altered sites is important for the RCA to be effective. The cost of monitoring reference sites can be substantial. These costs will vary depending on the regional specificity of the area. It is important to evaluate the trade-offs among allocating money for monitoring new reference sites, repeating reference sites, and impact sites. Increasing the number of reference sites might lead to a reduction in the frequency of re-evaluation of reference sites and/or a reduction in the number of impact sites considered for evaluation. A balance must be struck to optimize these sample size trade-offs.
3. Lower statistical power than a BACI design. Typically, an evaluation of the function of an impact site using the RCA will lead to more uncertainty than a well-designed BACI or BA approach. The RCA method trades off spatial specificity with statistical uncertainty. Developing a large reference site database, in most cases, will allow for comparisons to many impact sites. In contrast, a single BACI study will only be relevant to a single impact site, but will have much lower statistical uncertainty for detecting changes than the RCA.

Normal Range Approach

The Normal Range Approach applies the reference condition concept in that values measured at test sites are compared to some expected distribution of values derived from existing information on habitat requirements. The Normal Range Approach tests whether the values from the test site are outside the "normal" range of values. It is a flexible analytical framework that does not require a specific study design. The application of the Normal Range Approach typically involves a comparison of test (altered) and reference sites at the site and regional scale, but it can also be applied to data collected using other study designs, including Before-After and BACI designs, as well as the RCA (Kilgour et al. 2017).

The NRA uses information on habitat requirements or other measures of habitat or fish performance. Examples include changes in fish life histories as indicators of habitat change (Arciszewski et al. 2017) optimal values that may have been established from laboratory studies (e.g., dissolved oxygen, temperature); values that may exist in the literature, such as habitat preference data; or site-specific data collected for the question at hand. Models developed from field studies that describe relationships between habitat and measures of fish productivity (Sharma and Hilborn 2001) can also be used to determine benchmarks or threshold values for habitat indicators. This approach would be particularly useful for species that are well studied. Canada's Wild Salmon Policy (DFO 2005) has taken this approach to developing benchmarks from existing data sources for indicators of Pacific salmon habitat status (Stalberg et al. 2009).

A regional benchmarks framework has also been developed for evaluating fish productivity (Randall et al. 2017). The aim of this framework is to determine the lower threshold or benchmark for fish productivity associated with offset habitats. DFO (2014) also provides a framework that links fish productivity to habitat state, and can be used to develop benchmarks for indicator values that define the upper and lower bounds affecting productivity. Applying this framework for site-specific decision-making would require quantifying the productivity and state for the site in question, as well as threshold values that describe when changes in the habitat or species-states lead to relevant shifts in fish productivity; for example, when changes in a stressor no longer affect the productivity of a habitat or population, or conversely, when it compromises the ability of the habitat or species to contribute to the productivity of a fishery (DFO 2014). If these threshold values could be determined for specific management actions and stressors, and with sufficient information on the state and productivity of the ecosystem in question, the framework could provide a cost-effective approach to functional monitoring.

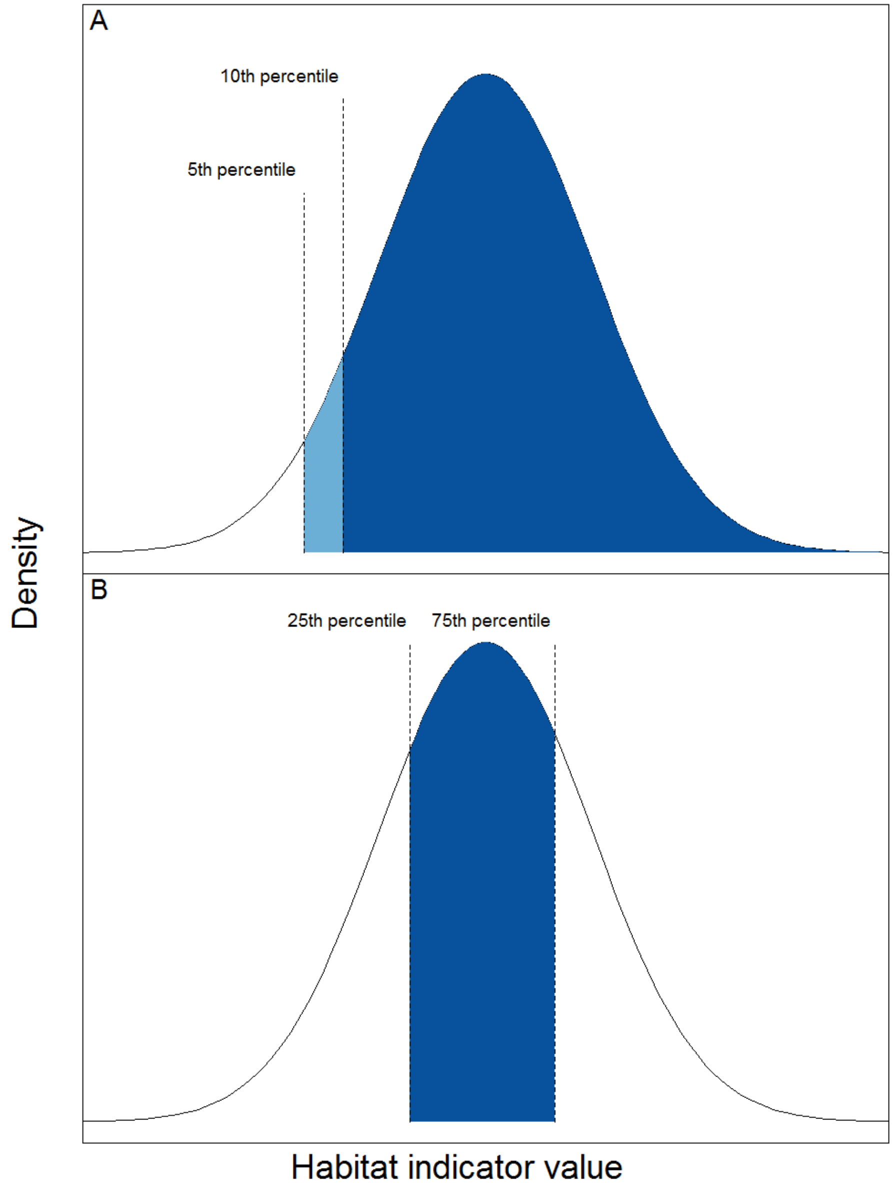


Figure 7. Hypothetical distribution of a habitat indicator value for a set of reference sites. A) If a habitat indicator value for an altered site fell within the shaded areas, the habitat indicator would be deemed to be within the normal range. The different levels of shading indicate the certainty with which a value is categorized as being within the normal range, and represents the Type I error rate; the light blue shading represents the 5th to 10th percentile of the distribution, and for values that fall within that region of the distribution there is less certainty that those values are in the normal range (i.e., unimpaired). The dark blue shading represents the 10th to 100th percentile of the distribution, and for values within that region of the distribution there is more certainty that those values are in reference condition. We illustrate how this would be used with the habitat indicator '% cover'. Typically, the amount of cover in a stream has a positive relationship with fish productivity; therefore, if the values for an altered site fell on the lower end of the distribution (e.g., 2% of the area assessed was cover), while 95% of the values for reference sites are > than 10%, it would be assumed, based on that indicator, that the stream is not in the normal range and is likely impaired. B) The shading indicates the values between the low (25th percentile - left) and high (75th percentile - right) benchmarks used by Anlauf et al. (2009). This example is more suited to indicators where there is an optimal value, such as spawning substrate size for salmonids, whereby too much small substrate will suffocate incubating eggs and too large substrate makes it difficult for spawning salmon to dig nests.

For physical habitat assessments, altered sites and reference sites are surveyed, and indicators are measured or scored. The altered site and reference site scores are compared to determine how close an altered site is to the normal range (Barbour 1998) (Figure 7). The distribution of habitat indicator values from reference sites are used to determine benchmarks that define the boundaries of different categories of impairment or function (Anlauf et al. 2009) (Figure 7B). For example, Anlauf et al. (2009) compares the values of 16 physical habitat parameters in coho salmon streams in Oregon between altered sites and reference sites. They use the 25th and 75th percentiles (representing low and high values, respectively) from the reference sites distribution to identify low < 25th%, moderate (> 25% and < 75%), and high (> 75%) benchmarks. These benchmarks can be used to determine where the values from altered sites fall with respect to reference sites, and thus the level of impairment or function. Error rates can be altered by selecting different decision rules, whereby lower and higher percentiles (e.g., 5th and 95th) will generally decrease the Type I error rate but increase the Type II error rates. How different decision rules influence error rates will depend on where the altered site values lie relative to the normal range (i.e., reference condition distribution).

Improvements to conventional methods for determining whether an altered site is outside the normal range are presented in Kilgour et al. (2017). These include ANOVA-based variants, and tolerance intervals. Tolerance intervals can be estimated for the upper and lower bounds of a reference distribution (e.g., 5th percentile of the distribution) and have the added benefit of including the uncertainty around the percentiles of the distribution.

Scores for habitat indicators based on qualitative assessment or expert opinion can also be used to evaluate habitat, where the sum of the scores represents the overall score for impairment or function (see Indices and Scores in the Measuring Function section). The score of an altered site is compared to scores from reference sites to determine its level of impairment (Barbour 1998). This is similar to developing benchmark values for different classifications of impairment and, based on our review, has received much less attention than the statistical methods applied to the Reference Condition Approach, which uses biological assemblage data.

For the approaches that apply the reference condition concept, the majority of monitoring effort lies in the development and maintenance of reference sites, or establishment of normal ranges or benchmarks, with much less effort spent on the evaluation of impact sites associated with each authorization (Barbour 1998). A large number of reference sites within a region are required to accurately determine where the impacted site fits relative to the reference condition. This asymmetry of effort and ongoing monitoring will be a major consideration in the development of a monitoring program that uses benchmarks or reference conditions as comparators for altered sites; however, costs can be alleviated by using existing information to develop benchmarks or reference conditions. Rather than each project being an independent assessment, the comparators require ongoing management and funding for monitoring to be effective. However, the costs of the program can be reduced by selecting metrics that are quick to measure, allowing for many sites to be sampled in a short period of time.

CONSIDERATIONS FOR ALL MONITORING DESIGNS

Statistical power

Determining the appropriate level of replication, whether it is the number of sites, years, or samples within a site, can be difficult and will vary among projects and indicators. Power analyses can be used to provide information about how replication will influence statistical power (i.e., the ability to detect an effect, if one exists). Statistical power is influenced by the spatial and temporal variation in the indicator (e.g., variance), the strength of the relationship between the indicator and the response (e.g., effect size), and the replication (e.g., number of sites, years, plots). Power analyses can be done using standard methods for simple statistical analyses, or simulations for more complex modeling approaches (Johnson et al. 2015). However, for such analyses, a number of assumptions about the spatial and temporal variability of indicators and effect sizes specific to the study and indicators are required. Assumptions can be refined using pilot studies and/or using values from the literature where comparable studies have been conducted. Power analyses should be used in the designing and planning stages of a monitoring program. General principles that can provide guidance when considering the different elements of replication (e.g., sites, years, plots, balance of the design) are outlined below.

Generally, increasing the number of years of data will increase statistical power. As more years of data are collected, the estimated means for an indicator approach the true means, reducing the likelihood of a spurious result. However, adding more years of data in an unbalanced design will have less of a benefit than if years are added to a balanced design. A balanced design where there are equal numbers of before and after years will provide more statistical power than an unbalanced design with the same total number of years. For example, O'Neal et al. (2016) compare the effectiveness of different types of restoration for salmonids. In their study they examined the effects of stream restoration activities on freshwater indicators by calculating the variance of the mean effects (i.e., impact – control) for different combinations of the numbers of years before and after restoration. Balanced designs always provided the lowest variance multiplier for a given number of years. They also show that increasing the number of years of data collected before restoration from one to two years had a greater decrease in the variance of their mean effect than increasing the number of years of data collected after restoration from one to 10 (Figure 8); this suggests that a minimum of two years of before-project data should be collected and matched with the collection of two years of after-project data. Acquiring data before a project can be challenging, but the trade-off between the number of years before and after, and the balance of the design make a strong case for investing in before-project data when possible. Since adding years to the before-project data collection is not possible after the alteration occurs, this approach requires that the design be determined in advance, and advocates pre-planning.

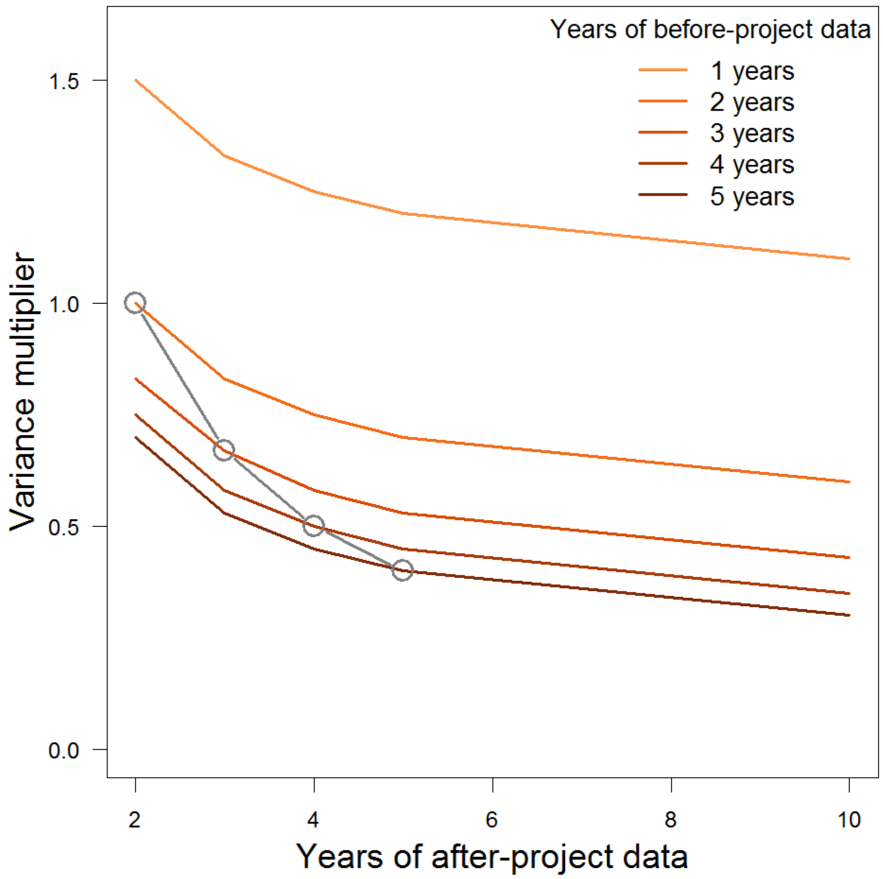


Figure 8. Plot of variance multipliers (relative variance) and the number of years of data collection before (coloured lines) and after a project was completed (x-axis). The variance multiplier represents the variance in the mean difference metric (impact – control). The grey lines and points indicate a balanced design where there are equal numbers of years of data collected before and after the project. Data were taken from O'Neal et al. (2016).

It is also important to consider the frequency of, and intervals between, sampling events. Using a 9-year study aimed at determining the effects of hydro peaking on benthic macroinvertebrate and fish communities, Smokorowski and Randall (2017) manipulated the temporal replication of a BACI design and asked if and how their results might change. They found that intermittent sampling (e.g., 1, 3, and 6 years after impact) provided uncertain results compared to continuous monitoring (e.g., 1, 2, 3, 4, 5, and 6 years after impact).

Generally, increasing the number of sites and samples within sites will increase statistical power. Beneditti-Cecchi (2001) used a simulation approach to examine how replication at multiple scales (i.e., number of sites, sampling events, and plots) influenced statistical power and Type I errors using a Beyond-BACI design consisting of up to four sites, eight sampling events (e.g., four years before and four years after the impact), and 10 replicate plots within a site. Statistical power was insensitive to how the number of sites and sampling events were allocated. For example, there was little difference in power if two sites were sampled four times compared to if four sites were sampled two times. This suggests that the variation among years and among sites were similar. However, there were substantial changes in power by increasing the number of plots sampled within a site. Across all design combinations of the number of sites and times sampled, the largest increase in power occurred when increasing replicates from one to two, and increases in power were often marginal after five plots had been sampled; this suggests that two to five replicate plots is the most cost-effective level of within-site replication, in this instance. Therefore, the minimum number of samples or replicates should be two but, when possible, a power analysis should inform the number of replicate samples.

Indicators with low spatial and temporal variability will require fewer samples to achieve the same statistical power as an indicator with high variability. Indicators can vary on many scales, and where or when the variability exists is important context for when it might be useful. For example, indicators such as water temperature often exhibit strong diurnal and seasonal variation. This variation can be reduced by measuring water temperature at a consistent time of year and day. Biological indicators (e.g., fish, invertebrate abundance and diversity) often exhibit high variation, regardless of when they are sampled, when compared to more static habitat indicators (e.g., channel morphology, cover). The distribution of variability across spatial and temporal scales can be estimated using mixed-effects models, and can provide valuable information about coordinating sampling effort (Li et al. 2001, Nakagawa et al. 2017).

The power to detect an effect, if it exists, is directly related to the size of the effect of interest. Functional monitoring may be appropriate for situations where project impacts are predicted to be small, readily mitigated, and/or are well understood (DFO 2012); therefore for some projects the effect size will be small. The low sampling effort of a functional monitoring program coupled with smaller effects sizes will make it difficult to detect changes in habitat function if they exist (Peterman 1990, Rubin et al. 2017). However, selection of simple metrics that exhibit low natural variability (e.g., substrate) may improve the ability to detect change; even more so than an intensive effectiveness monitoring program, which uses metrics with high natural variability (e.g., fish abundance). Matching the statistical design to the chosen metrics will be important for effective monitoring.

Weight of Evidence Approach

A Weight of Evidence Approach is a useful way to evaluate environmental change or impacts when data are limited and there may be multiple lines of evidence to evaluate. These approaches can be powerful in that they provide an informal but consistent way to combine all available data and/or indicators, including circumstantial (e.g., was there mitigation, data quantity) and quantitative data (e.g., changes in abundance, uncertainty of effects) in the impact assessment. As an example, Connors et al. (2014) developed a Weight of Evidence Approach through peer-review to evaluate the likelihood of impact from run-of-river hydro projects on salmonid populations. This was a particularly useful structure for these evaluations because the number of years of monitoring data and run-of-river project design varied among sites. Both factors are important for evaluating the effects on salmon populations since more years will provide greater statistical power and some project designs will be expected to have larger impacts on salmon over others. Given that functional monitoring is likely to have low power, and be applied to impacts that are well understood, this may be a useful approach. This could reduce likelihood of false positives triggering management actions, because results from monitoring would have to be considered alongside several other criteria that would be related to changes in habitat function as a result of current management measures.

No baseline data

The collection of baseline data requires that sufficient time and effort is expended prior to the development activity. Eighty-two percent of studies (63/77) we reviewed did not collect baseline data. In all cases where offset habitats are used as compensation, there will be no baseline data for the newly created habitat. However, in many cases offsets are in-kind habitat (new habitat is meant to have the same function as the impacted habitat) (Harper and Quigley 2005), which allows for comparisons between the new offset and pre-data from the old impacted habitats. When it is possible to collect data before and after a development activity, habitat data would be sufficient for a Before-After comparison in many instances, however this may not be the case for fish and other biotic indicators that are often highly variable in time and space. It may be unreasonable for some projects to wait for before-data to be collected prior to construction (e.g., bank stabilization for flood control). While there are monitoring designs that can be used to react to an immediate monitoring need (e.g., Control-Impact, Time-by-Trend Analysis, and the Normal Range Approach), they are often less desirable for the obvious reason that the comparison sites in most cases are confounded by unaccounted spatial and temporal variation.

Statistics and management decisions

Linking statistics to management objectives to inform decisions can be difficult. While some of the examples discussed in previous sections use classic null hypothesis testing and rely on the arbitrary significance (p-value < 0.05) of parameters to determine if there was an effect, this may not be the most appropriate approach to statistical inference, nor does it adequately communicate the uncertainty in specific parameter estimates. Null hypothesis testing is designed to determine how likely the alternative hypothesis (there was a change in habitat function) is if the null was true (no change in habitat function). This does not directly test if the alternative hypothesis is true (Bradford et al. 2017).

In many of the studies we reviewed, there was little to no discussion of whether a management measure was deemed successful in meeting management objectives. In contrast, O'Neal et al. (2016) evaluated 65 restoration projects, focusing on restoration of streams for improving salmon habitat in the Pacific Northwest, and presented the management decision criteria for determining project success. Their success criteria were well defined and transparent. They included the statistical test (e.g., t-test), the comparators used (e.g., before-project mean against after-project mean), the significance value required (e.g., p-value < 0.10), and the effect size required (e.g., 20% increase over baseline). Simple ways of communicating the effects of a test on a site or the performance of a management measure are valuable and help connect science to management. Effect sizes with measures of uncertainty (e.g., confidence limits) more readily communicate uncertainty and take full advantage of the data (Bradford et al. 2005) than simply reporting p-values.

It is also useful for managers to identify acceptable levels of uncertainty when making decisions. Bradford et al. (2017) suggest using an approach that defines acceptable levels of uncertainty by identifying a "range of indifference" when evaluating a change of a given magnitude. This tolerance for uncertainty or ranges of intolerance can be used to trigger or temper decisions (Bradford et al. 2017). For example, a tolerance rate of ± 20% change in an indicator, measured as the difference in the value from before-construction to after-construction, provides bounds for the level of change acceptable to a manager. This means, for a mean estimate of change of < 20% a manager would conclude the habitat remained the same and no further monitoring would be conducted, whereas if a change of >20% was observed additional monitoring would be required. This example only considers the mean change (effect size). This example can be taken a step further by considering the uncertainty in the mean change (Figure 9). Site 1 has a small and relatively certain change (mean = 5%, ± 7.5%), whereas it was estimated that for Site 2 the change was 17% with confidence intervals of ± 2%. In both cases, no further monitoring would be required and no further management actions would be taken. For Site 3, the mean effect was 12% - between Sites 1 and 2 - but the confidence intervals were ± 15% (i.e., confidence interval crosses 20%), which would have triggered a decision (e.g., continue functional monitoring, effectiveness monitoring, or additional compensation). This approach incorporates effect sizes and uncertainty into the decision process.

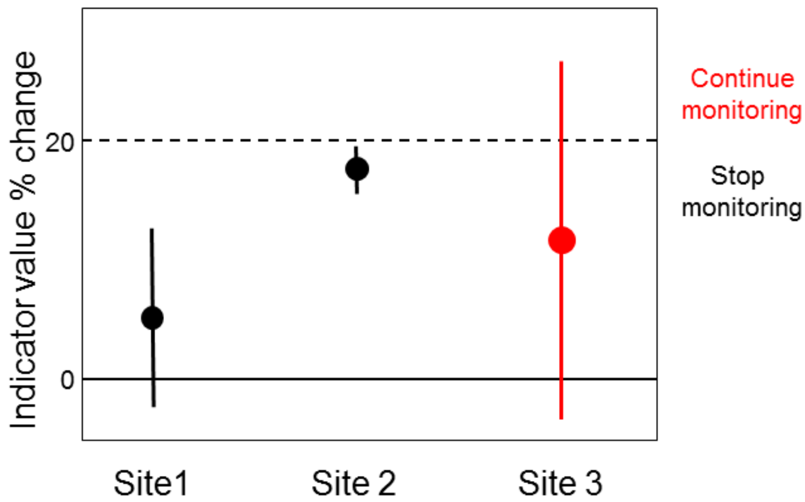


Figure 9. Hypothetical mean change and uncertainty (95% confidence intervals (CI)) in an indicator value for three test sites. The solid horizontal line represents no change, and dashed horizontal line represents the management relevant benchmark of 20% change. If a change exceeds this benchmark, monitoring will continue, whereas if the change is below this level, monitoring will stop. There was a 5% (lower CI: -2.5%, upper CI: 12.5%), 17% (lower CI: 15%, upper CI: 19%), and 12% (lower CI: -3%, upper CI: 27%), change at sites 1, 2, and 3. If a manager only considered the mean effect, no additional management activities would be triggered (i.e., monitoring would stop). However if the uncertainty around the mean change was considered, additional monitoring would be triggered. This example demonstrates how considering uncertainty could influence management decisions.

Management decisions that are based on statistical analyses that assess environmental change must consider not only the effect size but account for and quantify uncertainty. Furthermore, the management decision structure needs to incorporate this uncertainty into decision making. Not all analyses, however, provide effect sizes, incorporate or quantify uncertainty. Multivariate analyses that are currently used by the Canadian Aquatic Benthic Invertebrate Network, for example, do not provide effect sizes making it difficult to compare the effects across multiple studies as would be required for a meta-analysis. The current analytical methods used are unable to incorporate temporal replication and do not provide uncertainty for the estimates of site values in ordination space. To deal with this, simulations may be used to quantify effect sizes and some of the uncertainty, but these are somewhat disconnected from the original analysis and are not part of the standardized approach. Therefore, caution should be taken when using these monitoring frameworks to inform management decisions.

It will be important for managers to determine how the information from the functional monitoring programs will be interpreted and used to make decisions. Identifying what lines of evidence are required to make decisions and the decision outcomes available to managers, are critical steps that are often overlooked when designing a monitoring program, and need to be considered at the outset.

RAPID ASSESSMENT

An important feature of functional monitoring is that it often employs rapid assessment techniques. Rapid assessments are used around the world to determine the function of aquatic habitats (e.g., the United States, Australia, New Zealand, South Africa, European Union, and Canada). There are many definitions for what constitutes a “rapid” assessment in the literature. Definitions vary in how much of the assessment process is considered in determining if an assessment is rapid (e.g., field planning, data collection, laboratory analysis, data management, analysis, reporting). For example, the amount of time required to select sites (e.g., control or reference sites) or process samples and analyze data is often not considered, and can be substantial. Many studies use the definition provided by Sutula et al. (2006) and Fennessy et al. (2007), which define a rapid assessment as < 1 day to complete the monitoring (e.g., < 0.5 days for two people to conduct the field sampling and < 0.5 for two people for data management, analysis, and reporting). Cohen et al. (2005) defined it as < 1 hour per site. For the purposes of this report, we use the Sutula et al. (2006) and Fennessy et al. (2007) definitions of rapid assessment.

Studies were categorized as either rapid (based on the authors claim) or not (if not determined by the author we made our own assessment), or as unclear (if there was not enough information presented in the paper for us to assess). Forty percent of the studies were either not rapid (22%), according to our definition, or unclear (18%).

All physical habitat surveys that were declared rapid by the authors met our definition of rapid. However, 72% (28/39) of the studies that were declared rapid by the authors included macroinvertebrate metrics, meaning only the field component was actually rapid; thus, these studies were not considered rapid for our purposes. These studies were often associated with the term “rapid bioassessments”, which refers only to the field sampling. This is problematic because the effort (i.e., time and expertise) to identify the animals to an ecologically meaningful level of taxonomic resolution can be extensive. For instance, it has been estimated that it takes between 36 and 48 hours to count and identify all invertebrates from a single sample unit, of which multiple units are often collected (Pallottini et al. 2017). Fennessey et al. (2007) reviewed 40 rapid assessment methods for wetlands and was of the opinion that many methods were not rapid (i.e., < 1 day total - < 0.5 days in the field and < 0.5 days for analysis and reporting). However, there are some protocols that rely on low taxonomic resolution (e.g., order and some key families) and may fit in the more typical definition of rapid assessment (e.g., Parsons et al. 2002, Törnblom et al. 2011, Doll et al. 2016a).

None of the rapid assessments appeared to have considered field planning or site selection. While the time required to identify appropriate control sites, or to travel to field sites will be site specific and highly variable, these steps are common to all projects and will likely add substantial time to each field assessment. The time and cost of accessing a remote field site could be greater than the cost of conducting the survey itself. There was also no mention of using data loggers to monitor systems in our literature review. Data loggers can generate large amounts of useful data and require very little time in the field compared to field surveys. Although data loggers can be quick to deploy and retrieve, they often require more than one site visit, and in some instances, calibration (e.g., water depth loggers and discharge curves). Regardless, data loggers can be a cost-effective monitoring tool that should be considered for monitoring programs that include rapid assessment methods.

METRICS TO MEASURE FUNCTION

Ecological processes are complex, making them difficult to measure and quantitatively represent. Often, indicators measured in the field will only partially represent a component of the ecological or biological process of interest, and/or will be confounded by other processes that cannot be disentangled. The sensitivity of an indicator to changes in a process of interest (i.e., ecological change) is only one feature of many that are important to consider when selecting indicators for a monitoring program. Indicators can also vary in specificity, response time, natural variability, ease of measurement, and cost. Although this section provides a list of example metrics that could be measured, it will focus on approaches to measuring indicators and important features that should be considered when designing a functional monitoring program. Because functional monitoring is intended to be a scaled-down version of effectiveness monitoring, we will draw from effectiveness monitoring (Smokorowski et al. 2015) where applicable; other metrics that appear to be useful are drawn from the literature.

We categorized aquatic habitat by the function it provides to fish. Functional habitat monitoring is the assessment of whether management measures provide expected conditions suitable for fish to carry out their life processes. In the *Fisheries Act* these are identified as spawning ground, nursery, rearing, food supply, and migration areas. Appendix D presents lists of example indicators that would be suitable for functional monitoring for the five habitat types identified in the *Fisheries Act.* Measuring habitat function differs from physical habitat measurements in that measures of function are ecologically or biologically relevant to fish. Measures of habitat function could include physical measurements, but only if the proper context is provided. For example, categorizing substrate size based on the Wentworth scale (Wentworth 1922) is not particularly relevant classification for characterizing substrate size of fish spawning habitat because of the course scale and non-biological classification of size. Alternatively, quantifying or visually estimating the proportion of the substrate that is within a species' optimal size range for spawning and incubation can provide useful information for evaluating hypotheses about relationships between habitat and fish productivity (Braun and Reynolds 2014). The assumption that the optimal substrate size range will increase spawning success and the survival of embryos relates the physical measurement of substrate to spawning habitat function. Therefore, functional monitoring focuses on indirect measures of fish productivity, mainly physical and biological habitat features that are hypothesized to support fish production.

Broadly, there are four main classes of variable measurements that can be used to characterize the function of fish habitat: 1) indicators and surrogates, 2) indices and scores, 3) qualitative visual assessments using expert opinion, and 4) digital image assessments. We discuss these different types of measurements and their ability to detect change, effort, and ease of use.

INDICATORS AND SURROGATES

Indicators are quantities that describe changes in the state of another process, population, or habitat. In the context of functional monitoring of fish habitat, we define an indicator as some quantity that describes, and is hypothesized to be related to, changes in fish productivity. Indicators are more general quantities used to evaluate changes in fish productivity and are different from metrics, which are the specific representation or quantifications of an indicator; see Appendix D for examples. For context, temperature is an indicator and maximum temperature is a metric that is related to physiological stress and mortality in fish populations. Indicators may also be comprised of one or more quantitative metrics (e.g., multimetric index), or may be qualitative in nature ("loss of structure"). Indicators can provide multiple habitat functions. For instance, macrophyte beds provide both cover and food for juvenile fish, and the same metric can be used to represent both functions (e.g., shoot density and/or % macrophyte cover). However, other indicators that are related to multiple functions may require different metrics to describe the functions. Substrate can provide spawning and incubation habitat, as well as cover. Substrate metrics for spawning may quantify the percent of spawning substrate (amount of substrate within a specific size range), while cover may be the percent of substrate that is boulders. Ideally, indicators are measured in such a way that multiple metrics can be derived from one type of measurement, so that they can be used in assessing multiple habitat functions for various life-stages. A useful approach to capturing indicators that are highly variable or relevant to multiple life-stages (e.g., temperature and discharge; Appendix D) is the use of data loggers. Advancements in technology have made this a low cost option for monitoring over long time periods, for which a range of conditions can be captured and/or different metrics can be calculated.

Indices and scores

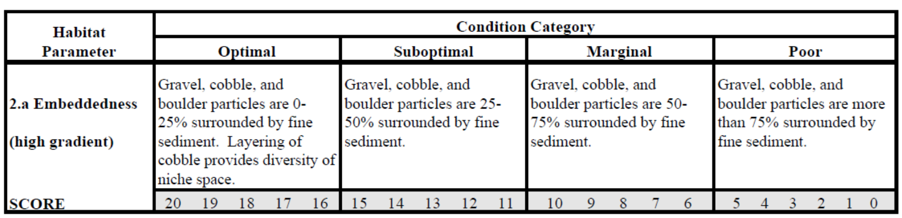
Indices and scores are types of indicators that are constructed by combining different sources of data that aim to provide a holistic representation of habitat status or function. Although indices can be seen as an attractive way to simplify complex ecological processes, address statistical issues associated with the number of variables in a model, and/or collinearity, there come with a number of challenges that warrant caution when using them for ecological monitoring (see Green and Chapman 2011 and references therein for a discussion).

Multimetric indices are often used when many metrics can be calculated from a single sample or survey, and metrics are correlated but represent slightly different components or processes of the ecosystem or habitat. For example, diversity metrics (evenness, richness, Shannon's diversity) are often combined using ordination methods such as Principal Components Analysis to generate a multimetric index of benthic macroinvertebrate assemblage. This approach has been used to develop sediment indices using sediment sensitive taxa (Turley et al. 2016). These indices may outperform single metrics in detecting differences in impact among sites (Talman et al. 1996) however, many have shown that derived indices lack stability in space and time making them highly specific to the conditions the data were collected under (Hamilton et al. 2010, Green and Chapman 2011). Some of the information contained in the raw data is also lost, which may lead to misleading interpretations of the data. Therefore, the development of indices must be driven by hypotheses and processes related to fish productivity and the data used to derive them should be presented to help alleviate the risk of misinterpretation. Indices are also used to deal with common statistical pitfalls, such as collinearity among variables that can reduce statistical power by inflating standard errors (Zuur et al. 2010). When variables are highly correlated it is simplest to use just one, whereas when there are many variables that are moderately correlated multivariate statistical methods such as ordination methods (e.g., principal components analysis, correspondence analyses and non-metric dimensional scaling) (Rice 2003) may be useful. Green and Chapman (2011) provide a critical assessment of indices and caution there use and highlight important considerations if they are used.

Habitat scores are another approach for combining multiple types of information, but differ from multimetric indices in that they are combining indicators (vegetation, macroinvertebrate, physical habitat) that aim to provide a habitat with a holistic score of impairment. There are a number of different ways to develop habitat function scores that range from simple to complex. The simplest approach is to qualitatively score habitat features during the assessment and then a sum of all the parameter scores is used to generate an overall habitat score (see Qualitative Visual Assessments using Expert Opinion section for more discussion). Others have used more complicated procedures that include scaling values between 0 and 1, averaging some scores, while multiplying others (Rowe et al. 2009). Regardless of the approach, all scores make assumptions about the weight of each parameter being measured in determining the status or function of habitat (Doll et al. 2016b). For example, giving equal weight to substrate and vegetation in an estuary assumes that these indicators are equally important to all fish species or fish assemblages, which may not be appropriate. Of the 40 rapid assessments for wetlands condition reviewed by Fennessey et al. (2007), many of the formulas and models used to calculate scores were considered unjustified by the authors. Habberfield et al. (2014) advocates for assessment scores to be developed based on ecological hypotheses. By using ecological hypotheses to develop assessment scores, the methods become transparent and justified. In addition, scores should be accompanied by the data used to generate them. This would present all of the information and highlight areas where the score might be misleading.

QUALITATIVE VISUAL ASSESSMENTS USING EXPERT OPINION

Qualitative assessments of habitat function are commonly used in rapid assessment protocols. These consist of data collection methods that are visual and subjective. This approach is popular because they serve as a quick and easy way to assess habitat. Qualitative assessments are commonly applied to assessments of cover (e.g., rating of 1-low cover to 5-high cover), substrate (e.g., substrate size classifications or quality), bank stability, and riparian vegetation and habitat classification (Doll et al. 2016b). These can be useful metrics but risk user bias and variability due to experience that can compromise the ability to compare among sites. Subjectivity in qualitative assessments can be reduced by providing details about the different classification or scoring methods and training. For example, Barbour (1998) provides detailed documentation about how to assess embeddedness of substrate that describes each condition category and provides pictorial examples (Figure 10). A detailed list of assumptions associated with visual assessments can increase transparency and reproducibility (Railsback and Kadvany 2008). Training and the use of simple visual survey methods, can also reduce observer subjectivity (Roper and Scarnecchia 1995). Testing the assessors can also achieve consistency by identifying where refinements need to be made to the protocol. The Province of British Columbia has developed a habitat watershed status evaluation protocol that focuses on fish values. This includes visual surveys for riparian and channel habitat, sediment transport, and fish passage (Pikard et al. 2014). Indicators are simple to assess and directly related to habitat function for fish. While their final status assessments are at the watershed level, the site level assessment protocols developed form the basis of the field monitoring and are applicable to functional monitoring. The trade-off between effort and information is important when considering qualitative visual assessments. Barbour (1998) cautioned using qualitative assessments because of their low precision and noted the additional value of quantitative assessments.



Habitat Parameter	Condition Category			
	Optimal	Suboptimal	Marginal	Poor
2.a Embeddedness (high gradient)	Gravel, cobble, and boulder particles are 0-25% surrounded by fine sediment. Layering of cobble provides diversity of niche space.	Gravel, cobble, and boulder particles are 25-50% surrounded by fine sediment.	Gravel, cobble, and boulder particles are 50-75% surrounded by fine sediment.	Gravel, cobble, and boulder particles are more than 75% surrounded by fine sediment.
SCORE	20 19 18 17 16	15 14 13 12 11	10 9 8 7 6	5 4 3 2 1 0

Figure 10. Description of condition categories for substrate embeddedness, taken from Barbour (1998).

DIGITAL IMAGE ASSESSMENTS

Digital images taken of benthic features can be a valuable approach to collecting data in the field. This method is being used for large-scale benthic monitoring programs, where species diversity is high with many species of algae, coral, vegetation, and animals that can be counted by applying a virtual quadrat (Perkins et al. 2016). This reduces field time, especially when using scuba, where bottom time is limited. Pictures have been used for stock assessment but we are unaware of any freshwater assessments using pictures to quantify species and/or habitat. This may be because of difficulty using this method in running waters, and the lack of species diversity in lake benthic environments. However, this remains a useful data collection approach.

Remote sensing is being rapidly developed and applied to aquatic environments (Marcus and Fonstad 2010). Remote sensing assessments of habitat could produce simple and fast assessments. For example, river channel morphology and instream habitat can be mapped using remote sensing (Legleiter et al. 2004), as well as macrophytes in lotic systems (Villa et al. 2017), estuaries, and near shore environments (Vahtmäe et al. 2006).

INDICATOR SELECTION

Selecting effective indicators can be difficult (Braun and Reynolds 2012) given the myriad of options that have been developed for measuring aquatic ecosystems. Although many aquatic indicators have been related, at some point, to measures of fish production, there are specific features that make some indicators more suitable for monitoring impacts on fish habitat and productivity. For indicators to be effective in functional monitoring programs, they must exhibit the following characteristics:

1. Indicators must be sensitive. Indicator values must change in relation to the management action; they must be measureable and occur in an appropriate time between the habitat alteration and measurement (e.g., < 6 months);
2. Indicators must not be hypersensitive. Indicators must not be more sensitive to other environmental changes not relevant to the management measure;
3. Indicators must exhibit low natural variability. Temporal and spatial variability must be low enough that the management measure signal is not masked by the noise of natural variability;
4. Indicators must be measured at the scale appropriate to the impact (e.g., landscape level measures may not detect point-source impacts);
5. There must be clear linkages between indicators and management objectives, which can be guided by conceptual model (e.g., Pathways of Effects);
6. There must be the ability to set reference points that will trigger management actions identified *a priori*. Without reference points or benchmarks it would be difficult to define when decisions would be triggered;
7. Indicator measurements, laboratory processing, and analyses must be rapid.

(adapted from (Rice 2003, Wieckowski et al. 2008))

Response time and the spatial scale will differ among indicators (Adams and Greeley 2000) (Figure 11). Basal indicators, such as water chemistry, will provide earlier responses than higher level indicators, such as organisms occupying higher trophic levels (e.g., macroinvertebrates, fish). Many evaluations of stream restoration, such as the addition of large wood, require multiple years before macroinvertebrate assemblages show a response (Entrekin et al. 2009). Indicators will also differ in the location or area they represent. For instance, two studies suggest that macroinvertebrate assemblages were more related to watershed and landscape characteristics than reach-scale physical habitat conditions (Miller et al. 2010, Louhi et al. 2011). These generalities are likely shared between marine and freshwater ecosystems.

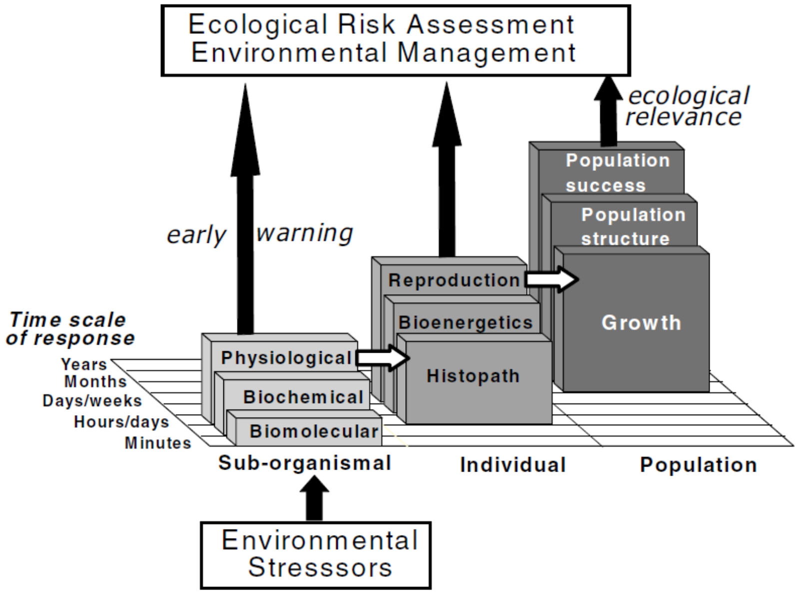


Figure 11. Taken from Adams and Greeley (2000), depicting the hierarchical responses of organisms to different stressors. Stressors observed through physiological, biochemical, or biomolecular responses are better early warning indictors, but may not be as ecologically relevant as higher-level population responses, such as indicators of population productivity.

Indicators may integrate the effects of ecological processes over larger scales (space and time), or provide information at finer scales that can be used to understand specific mechanisms of change (Rice 2003). For example, sediment loads in the substrate integrate the amount of sediment that is being discharged from the system over seasons, whereas suspended sediment may provide temporal information that is more useful for understanding mechanisms, because of its direct link to fish health and inter-gravel incubation conditions. Sampling contaminants in the environment over time will provide the conditions for a specific site and period. In contrast to direct measurements of contaminants, bioindicators can integrate environmental information over longer periods and larger spatial scales. For example, diatom assemblages can provide integrated information on environmental changes over multi-year periods (Summers et al. 2017), and contaminant loading in fish tissue can integrate over the fish's spatial range. Sessile species used as bioindicators, such as the Blue mussel (*Mytilus edulis*) (Beyer et al. 2017) and various macrophytes (e.g., *Zostera spp.*) (Farias et al. 2018) can indicate exposure to contaminants at a specific location, reducing the confounding effects of movement. Differences in tissue turnover rates can also integrate over different time scales; liver tissue will represent shorter periods compared to muscle, which will represent longer periods. The type of bioindicator will depend on the scale of the question in time and space.

Considering the scale an indicator represents is important when developing standardized monitoring programs because, while measuring many indicators with high specificity for a specific type of habitat change might make them more effective than indicators that integrate over larger scales, it will make it difficult to develop a standardized protocol. On the other hand, indicators that integrate do not provide a direct link to management actions. There is no perfect indicator, and the performance of indicators will vary across the criteria outlined above. It is important to recognize the trade-offs that exist between performance in one criteria compared to another. Therefore, the importance of each criterion should be considered in addition to indicator performance. While there are useful qualitative approaches to selecting indicators (see

Kershner et al. 2011), to date, there are no quantitative tools for assessing the performance of a proposed indicator.

Indicator selection and Pathways of Effects

The Fisheries Protection Program uses Pathways of Effects (POE) models to describe mechanisms by which an in-water activity (e.g., dredging, flow changes) affects fish productivity (DFO 2018b). Matching indicators to Pathways of Effects models is a useful conceptual framework for selecting indicators, and has been used to select risk-based indictors for marine ecosystems (O et al. 2015). Each indicator considered would be assigned to one or more pathways of effects depending on their specificity (Appendix E), and the metric measured would depend on the question being addressed by the data collected. Using the previous example, an indicator such as substrate would be assigned to multiple pathways, while the metrics used to describe the indicator will likely be specific to a particular pathway and endpoint. If substrate is used to indicate changes in cover, the percent boulders could be a metric; whereas if substrate was indicating food supply, embeddedness may be an appropriate metric. When applied to a project, the pathways of effects would first be determined, and then a project-specific monitoring plan (monitoring approach, site selection, and list of indicators to be measured) would be developed. Indicators would be checked off a comprehensive list (see Appendix E) and added to the monitoring plan based on the POEs associated with the project. This would provide consistency in the indicators measured for a given pathway, allowing for comparisons of management measures among projects addressing the same pathways of effects.

STANDARDIZED MONITORING

The development of a standardized monitoring program should be driven by the appropriate questions and a conceptual understanding of the systems it is being used to monitor, regardless of the region, system, or project application. Large-scale standardized monitoring programs often consist of a single monitoring design (e.g., RCA, reserve system) and several protocols that are adapted to different system types (e.g., wadeable streams, rivers, lakes, estuaries, wetlands); see Monitoring Design section for discussion of different monitoring designs and existing programs. Monitoring designs are general to all system types and it is possible to use one approach; however, some designs may be more effective than others, given the system, regional characteristics, and/or project. Similarly, protocols can be broadly applied to a system type, but many monitoring protocols are developed by states or regions, which suggest there may be benefits to the development of finer-scale protocols. Regional protocol development should not happen in isolation to achieve as much consistency as possible among comparable habitat types. Determining the scope of a standardized monitoring protocol is important because it is directly linked to its performance at achieving its objectives. A protocol that is too broad in scope will be ineffective; whereas the development of many protocols with high specificity will be expensive and will limit the comparisons among sites, thus reducing its power for evaluating the overall performance of different management measures. For example, finer-scale development of protocols may be useful for small streams, where there is high variability and many replicates in a small region. However, it may be difficult to develop regionally-based protocols for large estuaries (Imperial and Hennessey 1996) or rivers (Angradi et al. 2009). The level of replication for a system type and the monitoring design used should guide the scale at which protocols are developed.

Striking the appropriate balance with regards to the scope of standardized protocols is difficult, and the tipping point between a program that is too broad or specific will likely vary with region, habitat, species, and life stage. National oversight of the development of regional protocols will play an important role in meeting this challenge. For indicators that can be measured using a range of methods it will be important to ensure data are collected using a consistent method, so that data are comparable across protocols. For example, substrate composition can be collected using several methods that range in their precision and assumptions (e.g., sieved substrate samples, pebble count method, visual assessment of percent dominant substrate size). Ultimately, the objective may be the same for each method (e.g., represent the mean substrate size for a transect or area), but it may make comparisons among different methods difficult because each method has a different set of assumptions and biases. The sieved sample can be used to estimate substrate composition below the surface and will result in a mean substrate value with higher precision than the mean value derived from pebble counts of surface substrate. Furthermore, estimating a mean value from visual estimates of percent dominant substrate size is difficult and problematic for making comparisons with the other two methods, because a mean value cannot be calculated. Consistent methods at the national level, where possible, will avoid these challenges and increase the likelihood of meaningful comparisons when conducting meta-analyses of program performance.

Standardized monitoring programs provide several benefits. The benefits are often associated with the ease of implementation and monitoring, and include: 1) protocols to ensure metrics are measured correctly, 2) consistency in indicators measured and data entry, 3) determining the appropriate analyses prior to data collection, and 4) clear and standard reporting of results, with key information entered into an electronic database. Standardized monitoring can also be used by different agencies and/or private consultants to produce comparable data across a large number of sites and years that can be included in a national or regional database.

The value of a large national database that can be used to assess mitigation, offsetting, and restoration measures that are designed to reduce impacts to fish and fish habitat cannot be understated. Over time, as the number of projects assessed increases, powerful analytical methods such as meta-analyses can be used to determine how effective different management measures are at achieving sustainable fish productivity. Meta-analyses examine the overall effects of management measures by combining the results of several projects (Arnqvist and Wooster 1995). Meta-analyses are particularly useful when, at the project level, effects of a given management measure are small and uncertain, but when multiple project results are combined, they demonstrate a consistent pattern. This is particularly important in the context of functional monitoring because the uncertainty associated with individual studies/monitoring is high and inference will be weak. However, the collective power of these studies in a meta-analytical context will allow the performance of management measures to be evaluated. This information can then be used to adaptively manage the prescriptions associated with *Fisheries Act* authorizations and letters of advice. It is important to note that a meta-analysis has specific data requirements and standards; studies will be excluded from the analyses if they are of poor quality and/or inappropriate study design (e.g., no comparator). Therefore, it would be advantageous if all functional monitoring studies conformed to a minimum study design that would allow for its inclusion in a meta-analysis.

LIMITATIONS OF EXISTING PROGRAMS

The monitoring objectives of the large-scale agency-based programs that employ rapid bioassessments (many reviewed in the Monitoring Design section) may differ from that of functional monitoring to evaluate management measures. Specifically, most of the monitoring reported on large-scale environmental trends rather than the functional monitoring of small-scale activities, and therefore, the results from the review presented in this document should be taken accordingly. Bioassessment programs tend to focus on a broader scale of impacts and/or pollution, or water quality issues than the project-specific habitat alteration’s functional monitoring will be evaluating. Eighty-three percent (64/77) of the studies reviewed did not focus on the function of a specific habitat function (e.g., spawning, rearing, nursery), but rather the whole ecosystem, which was usually captured by metrics characterizing macroinvertebrate assemblages. It is important to note that the approaches and methods used in these large-scale monitoring programs may differ from the most appropriate approaches and methods for functional monitoring. Rubin et al. (2017) provided a critical review of the application of rapid bioassessment protocols to evaluating restoration activities in streams. They make two points relevant to the design of a functional monitoring program for evaluating management measures: 1) indicators typically used (i.e., macroinvertebrates) are often not justified, nor have clear relationships with how well the habitat functions for fish productivity; and 2) rapid assessments may not capture the appropriate scale of habitat alteration.

Rubin et al. (2017) suggest that while macroinvertebrates are good indicators of water quality and pollution, it is not clear if they respond to habitat alteration. They highlight the fact that the goal of restoration may not be to improve water quality but rather increase the quantity or quality of habitat for fish, and therefore the use of macroinvertebrates as indicators of reach-scale fish habitat function may not be appropriate. This point was also raised in a previous review of stream habitat restoration (Roni et al. 2008). Furthermore, commonly used metrics for macroinvertebrates, such as richness, diversity, or abundance are not clearly linked to changes in physical habitat and if changes in macroinvertebrate assemblages are observed, they do not always make sense. For example, a change in the abundance of macroinvertebrates at a particular site may be due to increases in a stress tolerant species or high nutrient inputs after logging, but may be incorrectly assessed as a positive response.

In their review, Rubin et al. (2017) go on to suggest that rapid assessment techniques are designed for examining large spatial extents (low sampling effort per site but many sites are sampled), but that there is little basis for using these approaches for assessing projects at the reach-scale (Rubin et al. 2017). As discussed earlier, it has been shown that macroinvertebrates are often more related to watershed characteristics than reach-scale characteristics (Miller et al. 2010, Louhi et al. 2011). Therefore, the prevalence of these rapid bioassessments based on macroinvertebrates found in the literature should not be misinterpreted as evidence of their effectiveness.

SUMMARY

Monitoring the effectiveness of management activities provides the Fisheries Protection Program information that can be used to adaptively manage at the project and program levels. Three levels of monitoring have been proposed to achieve the information needs of the FPP: 1) compliance, 2) functional monitoring, and 3) effectiveness monitoring. This document focuses on reviewing approaches and methods that could be used in a functional monitoring program. Below are a summary of some key points.

The conceptual understanding of the ecological system and processes affected by the management activities should drive the approaches and methods used in a functional monitoring program. Conceptual models such as the FPP’s Pathways of Effects models can be used to connect indicators to endpoints to monitoring program objectives. Explicitly connecting the lines between these different components is important for an effective monitoring program (Failing and Gregory 2003).

Monitoring the effectiveness of projects and the FPP requires a tiered approach, where the level of monitoring rigour matches the information needs of the FPP. In other words, functional monitoring, a scaled-down version of effectiveness monitoring, should be applied to projects where the impacts are relatively low and/or certainty is high (e.g., installation of a culvert in a stream during road building) compared to other projects where impacts may be high and/or uncertain (e.g., stream diversion that will alter habitat for a species at risk). The application of functional monitoring could also be structured based on previous monitoring results, similar to the tiered approach used by the Canadian Environmental Effects Monitoring Program (Hewitt et al. 2003), whereby functional monitoring is initially conducted to provide information that can be used to determine if more monitoring (i.e., effectiveness monitoring) is required.

Indicator selection can be difficult, given the myriad of potentially suitable options. The process of indicator selection should be driven by the monitoring objectives. Explicit thought about how an indicator will meet the monitoring objectives and achieve the management objectives is critical. Once this component of the indicator selection process has been met, other more quantitative assessments of indicator performance (e.g., sensitivity, response time, natural variability) can be used to further refine the selection of indicators to be monitored.

Identifying the appropriate balance between standardization and investigative rigour is a difficult but necessary task when designing a standardized monitoring program. Standardization facilitates broad application of the monitoring program but reduces the information specific to a project or management activity. Investigative rigour increases the value of the information that can be used to inform the effectiveness of management activities but increases time and costs of monitoring, which may be prohibitive for large-scale monitoring applications, and for making comparisons of management effectiveness across projects. Although a single standardized program is not possible given the diversity of habitats across Canada and the many management activities, it may be possible to design regionally-specific standardized functional monitoring programs that can be applied to projects with similar impact pathways, management activities, and site characteristics. The exact scale of standardization will ultimately depend on the information needs of the FPP.

REFERENCES CITED

Adams, S.M., and Greeley, M.S. 2000. Ecotoxicological indicators of water quality: Using multi-response indicators to assess the health of aquatic ecosystems. Environ. Sci. **123**: 103–115. doi:10.1023/a:1005217622959.

Angradi, T.R., Pearson, M.S., Jicha, T.M., Taylor, D.L., Bolgrien, D.W., Moffett, M.F., Blocksom, K.A., and Hill, B.H. 2009. Using stressor gradients to determine reference expectations for great river fish assemblages. Ecol. Indic. **9**(4): 748–764. doi:10.1016/j.ecolind.2008.09.007.

Anlauf, K.J., Jones, K.K., and Stein, C.H. 2009. The status and trend of physical habitat and rearing potential in coho bearing streams in the Oregon coastal coho evolutionary significant unit. OPSW-ODFW-2009-5.

Arciszewski, T.J., Munkittrick, K.R., Kilgour, B.W., Keith, H.M., Linehan, J.E., and McMaster, M.E. 2017. Increased size and relative abundance of migratory fishes observed near the Athabasca oil sands. Facets **2**(2): 833–858. doi:10.1139/facets-2017-0028.

Arnqvist, G., and Wooster, D. 1995. Meta-analysis: synthesizing research findings in ecology and evolution. Trends Ecol. Evol. **10**(6): 236–240. doi:10.1016/S0169-5347(00)89073-4.

Bailey, R.C., Linke, S., and Yates, A.G. 2014. Bioassessment of freshwater ecosystems using the Reference Condition Approach: comparing established and new methods with common data sets. Freshw. Sci. **33**(4): 1204–1211. doi:10.1086/678771.

Barbour, M.T. 1998. USEPA rapid bioassessment protocols for use in streams and wadeable rivers.

Benedetti-Cecchi, L. 2001. Beyond BACI: Optimization of environmental sampling designs through monitoring and simulation. Ecol. Apl. **11**(3): 783–799.

Beyer, J., Green, N.W., Brooks, S., Allan, I.J., Ruus, A., Gomes, T., Bråte, I.L.N., and Schøyen, M. 2017. Blue mussels (*Mytilus edulis spp.*) as sentinel organisms in coastal pollution monitoring: A review. Mar. Environ. Res. **130**: 338–365. doi:10.1016/j.marenvres.2017.07.024.

Bradford, M.J., Korman, J., and Higgins, P.S. 2005. Using confidence intervals to estimate the response of salmon populations (Oncorhynchus spp.) to experimental habitat alterations. Can. J. Fish. Aquat. Sci. **62**(12): 2716–2726. doi:10.1139/f05-179.

Bradford, M.J., Randall, R.G., Smokorowski, K.S., Keatley, B.E., and Clarke, K.D. 2014. A framework for assessing fisheries productivity for the Fisheries Protection Program. DFO. Can. Sci. Advis. Sec. Res. Doc. **2013/067** v + 25 p.

Bradford, M.J., Macdonald, J.S., and Levings, C.D. 2017. Monitoring fish habitat compensation in the Pacific region : lessons from the past 30 years. DFO. Can. Sci. Advis. Sec. Res. Doc. **2017/033** v + 26 p.

Braun, D.C., and Reynolds, J.D. 2012. Cost-effective variable selection in habitat surveys. Methods Ecol. Evol. **3**(2): 388–396. doi:10.1111/j.2041-210X.2011.00154.x.

Braun, D.C., and Reynolds, J.D. 2014. Life history and environmental influences on population dynamcis in sockeye salmon. Can. J. Fish. Aquat. Sci. **71**: 1198-1208.

Brinson, M.M., Rheinhardt, R., Applications, S.E., and Feb, N. 1996. The role of reference wetlands in functional assessment and mitigation. Ecol. Appl. **6**(1): 69–76.

Cohen, M.J., Lane, C.R., Reiss, K.C., Surdick, J.A., Bardi, E., and Brown, M.T. 2005. Vegetation based classification trees for rapid assessment of isolated wetland condition. Ecol. Indic. **5**(3): 189–206. doi:10.1016/j.ecolind.2005.01.002.

Connors, B.M. Marmorek, D.R., Olson, E., Hall, A.W., de la Cueva Bueno, P., Bensen, A., Bryan, K., Perrin, C., Parkinson, E., Abraham, D., Alexander, C., Murray, C., Smith, R., Grieg, L., Farrell, G. 2014. Independant review: Potential impacts of run-of-river hydroprojects on salmonids.

Crins, W.J., Gray, P.A., Uhlig, P.W.C., and Wester, M.C. 2009. The ecosystems of Ontario, Part 1: Ecozones and ecoregions. *In* Technical Report SIB TER IMA TR-01.

DFO. 2005. Canada's Policy for Conservation of Wild Pacific Salmon. *In* Wild Salmon Policy.

DFO. 2012. Assessing the effectiveness of fish habitat compensation in Canada: Monitoring design and metrics. DFO Can. Sci. Advis. Secr. Res. Doc. **2012/060**.

DFO. 2013. Fisheries Protection Policy Statement.

DFO. 2014. A science-based framework for assessing the response of fisheries productivity to state of species or habitats. Can. Sci. Advis. Secr. Sci. Advis. Rep. **2013/067**.

DFO. 2018a. Fisheries Protection Program. [accessed 1 January 2018].

DFO. 2018b. Pathways of Effects. [accessed 1 January 2018].

DFO. 2018c. Fisheries Act.

DFO. 2018d. An applicant's guide to submitting an application for authorization under Paragraph 35(2)(b) of the Fisheries Act. DFO.

Doll, B., Jennings, G., Spooner, J., Penrose, D., Usset, J., Blackwell, J., and Fernandez, M. 2016a. Identifying watershed, landscape, and engineering design factors that influence the biotic condition of restored streams. Water **8**(4): 1–18. doi:10.3390/w8040151.

Doll, B., Jennings, G., Spooner, J., Penrose, D., Usset, J., Blackwell, J., and Fernandez, M. 2016b. Can rapid assessments predict the biotic condition of restored streams? Water **8**(4): 1-22. doi:10.3390/w8040143.

Entrekin, S.A., Tank, J.L., Rosi-Marshall, E.J., Hoellein, T.J., and Lamberti, G.A. 2009. Response of secondary production by macroinvertebrates to large wood addition in three Michigan streams. Freshw. Biol. **54**(8): 1741–1758. doi:10.1111/j.1365-2427.2009.02223.x.

Evidence Collaboration for Environmental. 2018. Guidlines andstandards for evidence synthesis in environmental management.

Failing, L., and Gregory, R. 2003. Ten common mistakes in designing biodiversity indicators for forest policy. J. Environ. Manage. **68**(2): 121–132. doi:10.1016/S0301-4797(03)00014-8.

Farias, D.R., Hurd, C.L., Eriksen, R.S., and Macleod, C.K. 2018. Macrophytes as bioindicators of heavy metal pollution in estuarine and coastal environments. Mar. Pollut. Bull. **128**: 175–184. Elsevier. doi:10.1016/j.marpolbul.2018.01.023.

Feio, M.J., Viana-ferreira, C., and Costa, C. 2014. Testing a multiple machine learning tool (HYDRA) for the bioassessment of fresh waters. Freshw. Sci. **33**(4): 1286–1296. doi:10.1086/678768.

Fennessy, M.S., Jacobs, A.D., and Kentula, M.E. 2007. An evaluation of rapid methods for assessing the ecological condition of wetlands. Wetlands **27**(3): 543–560.

Green, R.H. 1979. Sampling design and statistical methods for environmental biologist. Wiley Interscience, Chichester, UK.

Green, R.H., and Chapman, P.M. 2011. The problem with indices. Mar. Pollut. Bull. **62**: 1377–1380. doi:10.1016/j.marpolbul.2011.02.016.

Habberfield, M.W., Blersch, S.S., Bennett, S.J., and Atkinson, J.F. 2014. Rapid geomorphic and habitat stream assessment techniques inform restoration differently based on levels of stream disturbance. J. Am. Water Resour. Assoc. **50**(4): 1051–1062. doi:10.1111/jawr.12156.

Hamilton, A.T., Stamp, J.D., and Bierwagen, B.G. 2010. Vulnerability of biological metrics and multimetric indices to effects of climate change. J. North Am. Benthol. Soc. **29**(4): 1379–1396. doi:10.1899/10-053.1.

Harper, D.J., and Quigley, J.T. 2005. No net loss of fish habitat: A review and analysis of habitat compensation in Canada. Environ. Manage. **36**(3): 343–355. doi:10.1007/s00267-004-0114-x.

Hawkins, C.P., Olson, J.R., and Hill, R.A. 2010. The reference condition: predicting benchmarks for ecological and water-quality assessments. J. North Am. Benthol. Soc. **29**(1): 312–343. doi:10.1899/09-092.1.

Hewitt, L.M., Dubé, M.G., Culp, J.M., MacLatchy, D.L., and Munkittrick, K.R. 2003. A proposed framework for investigation of cause for environmental effects monitoring. Hum. Ecol. Risk Assess. **9**(1): 195–211. doi:10.1080/713609859.

Imperial, M.T., and Hennessey, T.M. 1996. An ecosystem-based approach to managing estuaries: An assessment of the national estuary program. Coast. Manag. **24**(2): 115–139. doi:10.1080/08920759609362286.

Johnson, P.C.D., Barry, S.J.E., Ferguson, H.M., and Müller, P. 2015. Power analysis for generalized linear mixed models in ecology and evolution. Methods Ecol. Evol. **6**(2): 133–142. doi:10.1111/2041-210X.12306.

Kershner, J., Samhouri, J.F., James, C.A., and Levin, P.S. 2011. Selecting indicator portfolios for marine species and food webs: A Puget Sound case study. PLoS One **6**(10). doi:10.1371/journal.pone.0025248.

Kilgour, B.W., Somers, K.M., Barrett, T.J., Munkittrick, K.R., and Francis, A.P. 2017. Testing against "normal" with environmental data. Integr. Environ. Assess. Manag. **13**: 188-197.

Koning, C.W., Slaney, P.A., Gaboury, M.N., and Feduk, M.D. 1998. Techniques to evaluate the effectiveness of fish habitat restoration works in streams impacted by logging activities. Can. Water Resour. J. **23**(2): 191–203. doi:10.4296/cwrj2302191.

Legleiter, C.J., Roberts, D.A., Marcus, W.A., and Fonstad, M.A. 2004. Passive optical remote sensing of river channel morphology and in-stream habitat: Physical basis and feasibility. Remote Sens. Environ. **93**(4): 493–510. doi:10.1016/j.rse.2004.07.019.

Lewis, F.A., Harwood, A.J., Zyla, G., Ganshorn, K.D., and Hatfield, T. 2013. Long term aquatic monitoring protocols for new and upgraded hydroelectric projects. Can. Sci. Advis. Secr. Res. Doc. **2012/166**.

Li, J., Herlihy, A., Gerth, W., Kaufmann, P., Gregory, S., Urquhart, S., and Larsen, D.P. 2001. Variability in stream macroinvertebrates at multiple spatial scales. Freshw. Biol. **46**: 87–97. doi:10.1111/j.1365-2427.2001.00628.x.

Lindenmayer, D.B., and Likens, G.E. 2009. Adaptive monitoring: a new paradigm for long-term research and monitoring. Trends Ecol. Evol. **24**(9): 482–486. doi:10.1016/j.tree.2009.03.005.

Lindenmayer, D.B., and Likens, G.E. 2010. The science and application of ecological monitoring. Biol. Conserv. **143**(6): 1317–1328. doi:10.1016/j.biocon.2010.02.013.

Louhi, P., Mykrä, H., Paavola, R., Huusko, A., Vehanen, T., Mäki-Petäys, A., and Muotka, T. 2011. Twenty years of stream restoration in Finland little resonse by benthic macroinvertebrate communities. Ecol. Appl. **21**(6): 1950–1961. doi:10.1890/10-0591.1.

Marcus, W.A., and Fonstad, M.A. 2010. Remote sensing of rivers: The emergence of a subdiscipline in the river sciences. Earth Surf. Process. Landforms **35**(15): 1867–1872. doi:10.1002/esp.2094.

Miller, S.W., Budy, P., and Schmidt, J.C. 2010. Quantifying macroinvertebrate responses to in-stream habitat restoration: Applications of meta-analysis to river restoration. Restor. Ecol. **18**(1): 8–19. doi:10.1111/j.1526-100X.2009.00605.x.

Mudge, J.F., Barrett, T.J., Munkittrick, K.R., and Houlahan, J.E. 2012. Negative consequences of using $\alpha = 0.05$ for environmental monitoring decisions: A case study from a decade of Canada's environmental effects monitoring program. Environ. Sci. Technol. **46**(17): 9249–9255. doi:10.1021/es301320n.

Munkittrick, K.R. 2009. Ubiquitous criticisms of ecological field studies. Hum. Ecol. Risk Assess. **15**(4): 647–650. doi:10.1080/10807030903050616.

Nakagawa, S., Schielzeth, H., and Johnson, P.C.D. 2017. The coefficient of determination R^2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. J. R. Soc. Interface **14**(134): 1–35. doi:10.1098/rsif.2017.0213.

Neto, J.M., Barroso, D. V., and Barría, P. 2013. Seagrass quality index (SQI), a water framework directive compliant tool for the assessment of transitional and coastal intertidal areas. Ecol. Indic. **30**: 130–137. doi:10.1016/j.ecolind.2013.02.015.

O'Neal, J.S., Roni, P., Crawford, B., Ritchie, A., and Shelly, A. 2016. Comparing stream restoration project effectiveness using a programmatic evaluation of salmonid habitat and fish response. North Am. J. Fish. Manag. **36**(3): 681–703. doi:10.1080/02755947.2016.1165773.

O, M., Martone, R., Hannah, L., Greig, L., Boutillier, J., and Patton, S. 2015. An ecological risk assessment framework (ERAF) for ecosystem-based oceans management in the Pacific Region. DFO Can. Sci. Advis. Sec. Res. Doc. **2014/072** vii + 59 p.

Ogren, S.A., and Huckins, C.J. 2015. Culvert replacements: improvement of stream biotic integrity? Restor. Ecol. **23**(6): 821–828. doi:10.1111/rec.12250.

Pallottini, M., Goretti, E., Selvaggi, R., Cappelletti, D., Dedieu, N., and Céréghino, R. 2017. An efficient semi-quantitative macroinvertebrate multimetric index for the assessment of water and sediment contamination in streams. Inl. Waters **7**(3): 314–322. doi:10.1080/20442041.2017.1329912.

Parsons, M., Thoms, M., and Norris, R. 2002. Australian river assessment system: AusRivAS physical assessment protocol. Monit. River Heath Initiat. Tec. Rep. (22): 46.

Pearson, M.P., Quigley, J.T., Harper, D.J., and Galbraith, R.V. 2005. Monitoring and assessment of fish habitat compensation and stewardship projects: Study design , methodology and example case studies. Can. Manuscr. Rep. Fish. Aquat. Sci. **2729** xv + 124.

Perkins, N.R., Foster, S.D., Hill, N.A., and Barrett, N.S. 2016. Image subsampling and point scoring approaches for large-scale marine benthic monitoring programs. Estuar. Coast. Shelf Sci. **176**: 36–46. doi:10.1016/j.ecss.2016.04.005.

Perrin, C., Bennett, S., and Linke, S. 2007. Bioassessment of streams in north-central British Columbia using the reference condition approach. http://www.env.gov.bc.ca/epd/regions/skeena/water_quality/benthic/bio_streams_RCA_07.pdf.

Peterman, R.M. 1990. Statisitical power analysis can improve fisheries research and management. Can. J. Fish. Aquat. Sci. **47**: 2–15.

Pikard, D., Porter, M., Reese-Hansen, L., Thompson, R., Tripp, D., Carson, B., Tschaplinski, P., Larden, T., and Casley, S. 2014. Fish values: Watershed status evaluation, Version 1.0. Victoria, BC.

Railsback, S.F., and Kadvany, J. 2008. Demonstration flow assessment: Judgment and visual observation in instream flow studies. Fisheries **33**(5): 217–227. doi:10.1577/1548-8446-33.5.217.

Randall, R.G., Bradford, M.J., Clarke, K.D., and Rice, J.C. 2013. A science-based interpretation of ongoing productivity of commercial, recreational or Aboriginal fisheries. DFO. Can. Sci. Advis. Sec. Res. Doc. 2012/112 iv + 26 p.

Randall, R.G., Bradford, M.J., De Kerckhove, D.T., and Van Der Lee, A. 2017. Determining regional benchmarks of fish productivity using existing electrofishing data from rivers: proof of concept. DFO. Can. Sci. Advis. Sec. Res. Doc. **2017/018** v + 50 p.

Rice, J. 2003. Environmental health indicators. Ocean Coast. Manag. **46**(3–4): 235–259. doi:10.1016/S0964-5691(03)00006-1.

Robinson, D.L. 1987. Estimation and use of variance components. J. R. Stat. Soc. Ser. D. **36**(1): 3–14. doi:10.2307/2988267.

Roni, P., Hanson, K., and Beechie, T. 2008. Global review of the physical and biological effectiveness of stream habitat rehabilitation techniques. North Am. J. Fish. Manag. **28**(3): 856–890. doi:10.1577/M06-169.1.

Roper, B.B., and Scarnecchia, D.L. 1995. Observer variability in classifying habitat types in stream surveys. North Am. J. Fish. Manag. **15**(1): 49–53.

Rowe, D.K., Parkyn, S., Quinn, J., Collier, K., Hatton, C., Joy, M.K., Maxted, J., and Moore, S. 2009. A rapid method to score stream reaches based on the overall performance of their main ecological functions. Environ. Manage. **43**(6): 1287–1300. doi:10.1007/s00267-009-9302-z.

Rubin, Z., Kondolf, G.M., and Rios-Touma, B. 2017. Evaluating stream restoration projects: What do we learn from monitoring? Water **9**(3): 1–16. doi:10.3390/w9030174.

Schiff, R., Benoit, G., and MacBroom, J. 2011. Evaluating stream restoration: A case study from two partially developed 4th order Connecticut, U.S.A. streams and evaluation monitoring strategies. River Res. Appl. **27**(4): 431–460. doi:10.1002/rra.1365.

Sharma, R., and Hilborn, R. 2001. Empirical relationships between watershed characteristics and coho salmon (*Oncorhynchus kisutch*) smolt abundance in 14 western Washington streams. Can. J. Fish. Aquat. Sci. **58**(7): 1453–1463. doi:10.1139/f01-091.

Smokorowski, K.E., Bradford, M.J., Clarke, K.D., Clement, M., Gregory, R.S., and Randall, R.G. 2015. Assessing the effectiveness of habitat offset activities in Canada: Monitoring Design and Metrics. Can. Tech. Rep. Fish. Aquat.Sci. **3132:** vi + 48 p.

Smokorowski, K.E., and Randall, R.G. 2017. Cautions on using the Before-After-Control-Impact design in environmental effects monitoring programs. Facets **2**(1): 212–232. doi:10.1139/facets-2016-0058.

Stalberg, H.C., Lauzier, R.B., Macisaac, E.A., Porter, M., and Murray, C. 2009. Canada's Policy for Conservation of Wild Pacific Salmon: Stream, Lake, and Estuarine Habitat Indicators. Can. Manuscr. Rep. Fish. Aquat. Sci. **2859**: xiii + 135p.

Stewart-Oaten, A., and Bence, J.R. 2001. Temporal and spatial variation in environmental impact assessment. Ecol. Monogr. **71**(2): 305–339.

Stewart-Oaten, A., Murdoch, W.W., and Parker, K.R. 1986. Environmental impact assessment: "pseudoreplication" in time? Ecology **67**(4): 929–940.

Stoddard, J.L., Larsen, D.P., Hawkins, C.P., Johnson, R.K., and Norris, R.H. 2006. Setting expectations for the ecological conidition of streams: The concept of reference condition. Ecol. Appl. **16**(4): 1267–1276. doi:10.1890/1051-0761(2006)016.

Summers, J.C., Kurek, J., Rühland, K.M., Neville, E.E., and Smol, J.P. 2017. Assessment of multi-trophic changes in a shallow boreal lake simultaneously exposed to climate change and aerial deposition of contaminants from the Athabasca Oil Sands Region, Canada. Sci. Total Environ. **592**: 573–583. doi:10.1016/j.scitotenv.2017.03.079.

Sutula, M. a., Stein, E.D., Collins, J.N., Fetscher, a. E., and Clark, R. 2006. A practical guide for the development of a wetland assessment method: the California experience. J. Am. Water Resour. Assoc. **42**: 157–175. doi:10.1111/j.1752-1688.2006.tb03831.x.

Talman, S., O'Connor, N., Zampatti, B., and Cannan, F. 1996. Mount Lyell remediation: monitoring of benthic invertebrates in Macquaire Harbour, western Tasmania. *In* Department of Environment and Land Management. Department of the Environment and Energy. http://www.environment.gov.au/science/supervising-scientist/publications/ssr/monitoring-benthic-invertebrates-macquarie-harbour-western-tasmania.

Törnblom, J., Roberge, J.-M., and Angelstam, P. 2011. Rapid assessment of headwater stream macroinvertebrate diversity: an evaluation of surrogates across a land-use gradient. Fundam. Appl. Limnol. **178**(4): 287–300. doi:10.1127/1863-9135/2011/0178-0287.

Turley, M.D., Bilotta, G.S., Chadd, R.P., Extence, C.A., Brazier, R.E., Burnside, N.G., and Pickwell, A.G.G. 2016. A sediment-specific family-level biomonitoring tool to identify the impacts of fine sediment in temperate rivers and streams. Ecol. Indic. **70**: 151–165. doi:10.1016/j.ecolind.2016.05.040.

Underwood, A.J. 1991. Beyond BACI: Experimental designs for detecting human environmental impacts on temporal variations in natural populations. Mar. Freshw. Res. **42**(5): 569–587. doi:10.1071/MF9910569.

Underwood, A.J. 1994. On Beyond BACI : Sampling designs that might reliably detect environmental disturbances. Ecol. Appl. **4**(1): 3–15.

Vahtmäe, E., Kutser, T., Martin, G., and Kotta, J. 2006. Feasibility of hyperspectral remote sensing for mapping benthic macroalgal cover in turbid coastal waters - a Baltic Sea case study. Remote Sens. Environ. **101**(3): 342–351. doi:10.1016/j.rse.2006.01.009.

Villa, P., Pinardi, M., Toth, V.R., Hunter, P.D., Bolpagni, R., and Bresciani, M. 2017. Remote sensing of macrphytye morphological traits: Implications for the management of shallow lakes. J. Limnol. **76**: 109–126.

Vos, P., Meelis, E., and Ter Keurs, W.J. 2000. A framework for the design of ecological monitoring programs as a tool for enviornmental and nature management. Environmental Monit. Assess. **61**: 317–344. doi:10.1023/A:1006139412372.

Wentworth, C.K. 1922. A scale of grade and class terms for cladistic sediments. J. Geol. **30**: 377-392.

Wieckowski, K., Pickard, D., Porter, M., Robinson, D., Marmorek, D., and Schwarz, C. 2008. A conceptual framework for monitoring fisheries sensitive watersheds (FSW). Victoria.

Wiens, J.A., Parker, K.R. 1995. Analyzing the effects of accidental environmental impacts : Approaches and assumptions. Ecol. Appl. **5**(4): 1069–1083.

Zuur, A.F., Leno, E.N., and Elphick, C.S. 2010. A protocol for data exploration to avoid common statistical problems. Methods Ecol. Evol. **1**(1): 3–14. doi:10.1111/j.2041-210X.2009.00001.x.

GLOSSARY

Term	Description	Reference
BA, Before-After	A commonly used monitoring design that compares data collected before and after a development activity.	Underwood 1991
BACI, Before-After-Control-Impact	A commonly used monitoring design where the control and impact sites are sampled before and after the development occurs.	Underwood 1991
Bioassessment	An evaluation of the condition of a waterbody using biological samples and other direct measurements of the resident biota in surface water.	Barbour 1998
CI, Control-Impact	A commonly used monitoring design that compares data between control and impact sites.	Underwood 1991
Effectiveness Monitoring	A science-based activity, requiring a standardized, transferable design. The metrics or indicators must measure productive capacity or fish-based surrogates of productive capacity.	Smokorowski et al. 2015
Fish Productivity	A survival parameter specific to a population of fish (e.g., maximum growth rate of a population at low density). Productivity may also be characterized by other population traits such as growth, fecundity and age-at-maturity.	Randall et al. 2013
Fisheries Productivity	The sustained yield of all component populations and species, and their habitat, which support and contribute to a fishery in a specified area.	Randall et al. 2013
Functional Monitoring	A science-based, scaled-down version of effectiveness monitoring that relies on surrogate metrics to assess whether management measures provide expected conditions suitable for fish to carry out their life processes.	DFO 2012
Habitat	Spawning grounds and other areas, including nursery, rearing, food supply, and migration areas, on which fish depend directly or indirectly in order to carry out their life processes.	DFO 2018c
Indicator	Some quantity that describes, and is hypothesized to be related to, changes in fish productivity. Indicators may be comprised of one or more quantitative metrics, or may be qualitative in nature (cf. “change in LWD”, “loss of structure”).	Bradford et al. 2014

Term	Description	Reference
Letters of Advice	Guidance provided to a proponent by the FPP when a Fisheries Act Authorization is not required, but there is potential to avoid or mitigate any effects of the project impact on fisheries productivity.	DFO 2018d
***Fisheries Act* Authorizations**	Guidance for a project proponent from the FPP outlining how to avoid or mitigate impacts on fisheries productivity where possible, and requirements for restoration and offsetting where impacts are unavoidable and cannot be mitigated.	DFO 2018d
Management Monitoring Objectives	Monitoring objectives of the Fisheries Protection Program related to project monitoring are: 1) to ensure conformity with advice, construction/design standards and compliance with the *Fisheries Act* and *Species at Risk Act* (compliance monitoring program); and 2) to evaluate the effectiveness of management measures aimed at reducing the impacts of projects on fish and fish habitat (functional and effectiveness monitoring programs).	This Report
Measurements	Measurements are taken in the field and describe the current state of the ecosystem or its biota. Examples include fish abundance or discharge.	Bradford et al. 2014
Meta-analyses	A powerful analytical method that can be used to determine how effective different management measures are at achieving sustainable fish productivity by evaluating the overall effect of a given management measure for multiple projects.	Arnqvist and Wooster 1995
Metric	The specific representation or quantification of an indicator. Metrics are used to evaluate change or the relationship between the altered site and control(s) or relevant comparator(s). A metric can be derived from before-after field measurements (e.g., change in fish abundance), or can be estimated from baseline measurements and a predicted or modelled effect.	Bradford et al. 2014
Mitigation	Is a measure to reduce the spatial scale, duration, or intensity of serious harm to fish that cannot be completely avoided. Mitigation measures include the implementation of best management practices during the construction, maintenance, operation and decommissioning of a project.	DFO 2013

Term	Description	Reference
Multimetric Indices	A type of indicator that is constructed by combining different sources of data aimed at providing a holistic representation of habitat status or function.	This Report.
NRA, Normal Range Approach	An approach that compares a test site to the distributions of conditions defined by multiple reference sites that represent some desirable state (e.g. undisturbed, pristine, or not-impaired).	
Offsetting	A measure that counterbalances unavoidable serious harm to fish resulting from a project with the goal of maintaining or improving fish productivity.	DFO 2013
Program Monitoring Objectives	Monitoring objectives specific to each program with regards to how they measure the effectiveness of management measures. For example, functional monitoring will measure effectiveness by evaluating if the management measures are functioning as intended, whereas effectiveness monitoring will evaluate the productivity-based milestones.	This Report, DFO 2012
Project Monitoring Objectives	Monitoring objectives specific to each project with regards to how they measure the effectiveness of management measures. For example, the project objectives for evaluating a culvert may be to determine if the culvert provides upstream and downstream passage; whereas the evaluation of new spawning habitat for adult salmon may be to determine if the habitat provides cover and spawning gravel for adult salmon.	This Report
Quantitative	Collecting both physical and biological measures, metrics, and indicators through measurement to generate numerical data.	Bradford et al. 2014
Qualitative	Collecting both physical and biological measures, metrics, and indicators through descriptive assessments of the state of a feature, which may be unitless.	Bradford et al. 2014
Range of Indifference	An approach that identifies acceptable levels of uncertainty when making decisions. This tolerance for uncertainty, or ranges of intolerance, can be used to trigger or temper decisions.	Bradford et al. 2017

Term	Description	Reference
Rapid Assessment	An assessment protocol that can be conducted in a short amount of time (e.g., < 1 day for two people to collect the data, manage the data, analyze the data, and complete reporting).	Sutula et al. 2006
RCA, Reference Condition Approach	An approach that compares a test site to a set of conditions defined by multiple reference sites that represent some desirable state (e.g. undisturbed, pristine, or not-impaired).	Stoddard et al. 2006
Restoration	The creation or restoration of a previously degraded habitat known to have served a specific function in the past.	Smokorowski et al. 2015
Standardized Monitoring	Monitoring programs that use consistent data collection, analysis, and reporting protocols.	This Report.
System Type	Lake, river, stream, estuary, marine, coastal, or other major category of waterbody.	This Report.

APPENDIX A

Table A 1. Metadata from reviewed literature.

Reference Id	Paper ID Number
Reference	Authors and Year of Paper
Publisher	Journal or Report Series Name
Title	Title of the Paper
Location	Country/State or Province, Country/Country in alphabetical order
Ecosystem	Freshwater, Marine
Habitat Type	Estuary, Lake, Near-shore, Pond, River, Stream, Wetland
Habitat Function	Cover, Ecosystem Function, Food, Migration, Nursery, Rearing, Spawning
Biological Physical Response	This category describes the broad response examined in the paper and is defined by the paper. Terms or categories are grouped when necessary
Development Type	Defined by the paper but broad categories are used
Intervention	The type of intervention defined by the paper but similar techniques have been grouped
Data Type	The type of data collected for describing the response metrics
Monitoring Design	This is the design employed to determine if change has occurred (comparator)
Standardized Protocol	What standardized protocol was used, this has to be named in the paper
Before Monitoring (Years)	Number of years before the impact
After Monitoring (Years)	Number of Years after the Impact
Years Post Intervention	Number of Years Post Intervention
Number of Control Sites	Number of Control Sites
Number of Altered Sites	Number of Altered or Impacted Sites
Statistical Analyses	What analysis was used in the paper to test for a change in habitat/other types of modeling that are used to develop indices or scores
Rapid < 1 Day	Was the survey rapid, most papers don't call their surveys rapid and rapid can mean different things. This provides a standardized evaluation of the amount of effort

Table A 2. Extracted data from reviewed literature.

Reference ID	Reference	Publisher	Title	Location	Ecosystem	Habitat Type	Habitat Function	Biological Physical Response	Development Type	Intervention	Data Type	Monitoring Design	Standardized Protocol	Before Monitoring Years	After Monitoring Years	Years Post Intervention	Number of Control Sites	Number of Altered Sites	Statistical Analyses	Rapid < 1 day
1	Adomato et al. 1997	US Fish and Wildlife Service	The use of rapid bioassessment protocols to describe fish and benthic macroinvertebrate communities in three creeks near the Little River National Wildlife Refuge, Mccurtain County Oklahoma	United States/Oklahoma	Freshwater	Stream	Ecosystem Function	Fish Assemblage/Macroinvertebrates	Industrial Sites	None	Quantitative	CI	USEPA Rapid Bioassessment Protocols for Streams and Rivers	0	0	0	3	3	None	Yes
2	Albertson et al. 2013	River Research and Applications	How does restored habitat for chinook salmon (oncorhynchus tshawytscha) in the merced river in california compare with other chinook streams?	United States/California	Freshwater	Stream	Rearing	Salmon Density Per Discharge	Mine	Channel Reconstruction	Quantitative	RC	USEPA Rapid Bioassessment Protocols for Streams and Rivers	0	1	0	19	1	Compare Distributions/Correlation/PCA	Yes
3	Alford 2014	Hydrobiologia	Multi-scale assessment of habitat and stressors	United States/Louisiana	Freshwater	Stream	All	Fish Assemblage	Multiple Stressors	None	Mix	Random Sample	USEPA Rapid Bioassessment Protocols for Streams and Rivers	0	0	0	Not Described	50	Partial Redundnacy Analysis	Yes
4	Angradi et al. 2009	Environmental Monitoring and Asessment	Using stressor gradients to determine reference expectations for great river fish assemblages	United States	Freshwater	River	Ecosystem Function	Fish Assemblages/Physical Habitat/Water Chemistry	Multiple Stressors	Not Described	Quantitative	RC	Environmental Monitoring and Assessment Program for Great River Ecosystems	1	1	0	0	Many - Not described	Cumulative Density Function	No
5	Anton et al. 2011	Ecology of Freshwater Fish	Restoration of dead wood in Basque stream channels: Effects on brown trout population	European Union/Spain	Freshwater	Stream	Spawning/Rearing	Fish Density	Snagging	Large Wood Addition	Quantitative	BACI	None	2	2	0	4	4	Anova	Yes
6	Arthur and Kauss 2000	Ontario Minsistry of Environment	Sediment and Benthic Community Assessment of the St. Marys River	Canada/Ontario	Freshwater	River	Ecosystem Function	Macroinvertebrates/Water Chemistry	Industrial Sites	None	Quantitative	Random Sample	None	0	0	0	0	8	Correlation/Discriminant Function Analysis/PCA	Unclear
7	Benedetti-Cecchi and Osio 2007	MEPS	Replication and mitigation of effects of confounding variables in environmental impact assessment: Effect of marinas on rocky-shore assemblages	European Union/Italy	Marine	Near-shore	Ecosystem Function	Algae Assemblage/Macroinvertebrates	Marina	None	Quantitative	CI	None	0	1	40	9	2	Anova/Permanova	Yes
8	Bennett et al. 2009	Gitxsan Forest Enterpris	Bioassessment of streams in Northwest BC	Canada/British Columbia	Freshwater	Stream	Ecosystem Function	Macroinvertebrates	Forestry	None	Mix	RC	Skeena BEAST/USEPA Rapid	0	0	0	143	7	Discriminant Function Analysis/NMDS	Yes

Reference ID	Reference	Publisher	Title	Location	Ecosystem	Habitat Type	Habitat Function	Biological Physical Response	Development Type	Intervention	Data Type	Monitoring Design	Standardized Protocol	Before Monitoring Years	After Monitoring Years	Years Post Intervention	Number of Control Sites	Number of Altered Sites	Statistical Analyses	Rapid < 1 day
		es and Westfraser Mills Ltd.	using the Skeena BEAST09										Bioassessment Protocols for Streams and Rivers							
9	Berkowitz et al. 2013	Soil Science Society of America Journal	Linking Wetland Functional Rapid Assessment Models with Quantitative Hydrological and Biogeochemical Measurements across a Restoration Chronosequence	United States/Louisiana	Freshwater	Wetland	Ecosystem Function	A Horizon Biomass/Cation Exhange Capacity/Flood Frequency/Ground Vegetation Cover/O Horizon Biomass/Shrub-Sapling Density/Snag Density/Tree Basal Area/Woody Debris Biomass	Deforestation	None	Quantitative	RC	Hydrogeomorphic	0	1	1-20	21	45	Pearson's Correlation	Yes
10	Bonada et al. 2006	Journal of North American Benthological Society	A comparison of rapid bioassessment protocols used in 2 regions with Mediterranean climates, the Iberian Peninsula and South Africa	Iberian Peninsula/South Africa	Freshwater	Stream	Ecosystem Function	Macroinvertebrates	Multiple Stressors	None	Quantitative	RC	Iberian Penninsula Rapid Bioassessment/South African Rapid Bioassessment	0	0	0	7	4	None	Yes
11	Borisko et al. 2007	Water Quality Research Journal of Canada	An evaluation of rapid bioassessment protocols for stream benthic invertebrates in Southern Ontario, Canada	Canada/Ontario	Freshwater	Stream	Ecosystem Function	Macroinvertebrates	None	None	Quantitative	RC	CABIN/MNRF/OBBN Rapid Bioassessments/TRCA	0	0	0	0	11	Anova	Yes
12	Boys et al. 2012	Journal of Applied Ecology	Improved fish and crustacean passage in tidal creeks following floodgate remediation	Australia/New South Wales	Freshwater	River/Stream	Nursery	Crustacean Passage/Fish Passage	Agriculture	Flood Gate Control	Quantitative	BACI	None	0	1	0	8	3	Permanova	Yes
13	Brown et al. 2016	Science of the Total Environment	Macroinvertebrate community assembly in pools created during peatland restoration	United Kingdom	Freshwater	Wetland	Ecosystem Function	Macroinvertebrates/Physicochemical	Flow Alteration	Pool Addition	Quantitative	RC	None	0	0	5-10	20	20	NMDS/Null Models/Permanova	Yes
14	Cahill et al. 2015	North American Journal of Fisheries Management	Assessing Responses of Fish to Habitat Enhancement in Barrenlands Streams of the Northwest Territories	Canada/Northwest Territories	Freshwater	Stream	Migration	Fish Passage	Mine	Fishway	Quantitative	BACI	None	1	2	0	2	3	Chi-Square	No
15	Chamberlain and Brooks 2016	Ecological Indicators	Testing a rapid Floristic Quality Index on headwater wetlands in central Pennsylvania, USA	United States/Pennsylvania	Freshwater	Wetland	All	Floristic Assemblage	Gradient of Stress	None	Quantitative	RC	Rapid Fish Quality Index	0	0	0	0	87	Anova	Yes
16	Cianfrani et al. 2001	(to be submitted to the	Assessment of Urban Streams in Fairmount Park,	United States/Pennsylvania	Freshwater	Stream	Ecosystem Function	Geographic Condition/Physical Habitat/Riparian Condition/Macroinvertebrates	Urbanization	None	Mix	RC	USEPA Rapid Bioassessment Protocols for	0	0	0	16	426	None	Yes

Reference ID	Reference	Publisher	Title	Location	Ecosystem	Habitat Type	Habitat Function	Biological Physical Response	Development Type	Intervention	Data Type	Monitoring Design	Standardized Protocol	Before Monitoring Years	After Monitoring Years	Years Post Intervention	Number of Control Sites	Number of Altered Sites	Statistical Analyses	Rapid < 1 day
			Philadelphia, PA										Streams and Rivers							
17	Cohen et al. 2006	Journal of the American Water Resources Association	Vegetation based classification trees for rapid assessment of isolated wetland condition	United States/Florida	Freshwater	Wetland	All	Floristic Assemblage	Agriculture/Urbanization	None	Quantitative	RC	Wetland Condtion Index	0	0	0	73	120	Classification Trees/Regression Trees	Yes
18	Cooperman et al. 2007	Fisheries	Streambank Restoration Effectiveness : Lessons Learned from a Comparative Study	Canada/British Columbia	Freshwater	Stream	Spawning/Rearing	In-Channel Habitat/Macroinvertebrates/Riparian Vegetation	No Development	Bank Stabilization	Quantitative	RC	None	0	1	3-8	11	16	Anova	Yes
19	Cordell and Toft 2012	US Fish and Wildlife Service	2010 Invertebrate Monitoring at Duwamish Waterway Restoration Sites : Hamm Creek, Herring's House, Northwind's Weir, and Kenco Marine	United States/Washington	Marine	Estuary	Ecosystem Function	Benthic Meiofauna/Macroinvertebrates/Riparian Insects	Industrial Sites	Adding Off-Channel Features/Planting Emergent and Riparian Vegetation/Reduce Armoring	Quantitative	RC	None	0	0	4-11	2	4	Permanova	Yes
20	de Bikuna et al. 2015	Fundamental and Applied Limnology	Development of a multimetric benthic macroinvertebrate index for assessing the ecological condition of Basque streams (north of Spain)	European Union/Spain	Freshwater	Stream	Ecosystem Function	Macroinvertebrates	Urbanization	None	Quantitative	RC	None	0	0	0	7	15	Box-Wisker Plots/U-test	Unclear
21	de Mutsert and Cowan 2012	Estuaries and Coasts	A Before-After-Control-Impact Analysis of the Effects of a Mississippi River Freshwater Diversion on Estuarine Nekton in Louisiana, USA	United States/Louisiana	Marine	Estuary	Nursery	Nekton	Dyking	Freshwater Diversion	Quantitative	RC	None	1	14	0	6	13	MDS/Permanova	Unclear
22	Doll et al. 2015	Journal of the American Water Resources Association	Evaluating the eco-geomorphological condition of restored streams using visual assessment and macroinvertebrate metrics	United States/North Carolina	Freshwater	Stream	Ecosystem Function	Physical Habitat/Bedform/Macroinvertebrates/Mophologic Condition	Urbanization	Bank Stabilization	Mix	RC	NCSU Stream Performance Assessment	0	1	1-15	42	114	Multiple Regression/PCA	Yes
23	Doll et al. 2016	Water	Can rapid assessments predict the biotic condition of restored streams?	United States/North Carolina	Freshwater	Stream	Ecosystem Function	Physical Habitat/Bedform/Macroinvertebrates/Mophologic Condition	Urbanization	Bank Stabilization/Channel Modification/Large Wood Addition	Mix	RC	NCSU Eco-Geomopholgical Assessment/NCSU Stream Performance Assessment/Peterson's Riparian Channe and Environmental Inventory/USDA Stream Visual	0	1	1-15	0	65	Multiple Regression/PCA	Yes

Reference ID	Reference	Publisher	Title	Location	Ecosystem	Habitat Type	Habitat Function	Biological Physical Response	Development Type	Intervention	Data Type	Monitoring Design	Standardized Protocol	Before Monitoring Years	After Monitoring Years	Years Post Intervention	Number of Control Sites	Number of Altered Sites	Statistical Analyses	Rapid < 1 day
													Assessment Protocol/USEPA Rapid Bioassessment Protocols for Streams and Rivers							
24	Doll et al. 2016	Water	Identifying watershed, landscape, and engineering design factors that influence the biotic condition of restored streams	United States	Freshwater	Stream	Ecosystem Function	Landscape/Macroinvertebrates/Restoration Design	Multiple Stressors	Channel and Floodplain Reconstruciton/Large Wood Addition/Rock Addition	Quantitative	Targeted	EPT	0	0	1-10	0	79	PCA/Ridge Regression	Yes
25	Entrekin et al. 2009	Freshwater Biology	Response of secondary production by macroinvertebrates to large wood addition in three Michigan streams	United States/Michigan	Freshwater	Stream	Ecosystem Function	Macroinvertebrates	Deforestation	Large Wood Addition	Quantitative	CI	None	1	2	0	3	3	Anova	No
26	Flores et al. 2017	Ecological Engineering	Effects of wood addition on stream benthic invertebrates differed among seasons at both habitat and reach scales	European Union/Spain	Freshwater	Stream	Ecosystem Function	Macroinvertebrates/Physical Habitat	Deforestation	Large Wood Addition	Quantitative	BACI	None	1	1	0	4	4	Mixed-Effects/Permanova	Yes
27	Franklin and Bartels et al. 2012	Aquatic Conservation: Marine and Freshwater Ecosystems	Restoring connectivity for migratory native fish in a New Zealand stream: effectiveness of retrofitting a pipe culvert	New Zealand	Freshwater	Stream	Migration	Fish Assemblage/Fish Passage	Bridge	Retrofitting Culvert	Quantitative	BA	None	3	3	0	0	1	Anova	No
28	Goodman et al. 2015	Restoration Ecology	A mapping technique to evaluate age-0 salmon habitat response from restoration	United States/California	Freshwater	Stream	Rearing	Physical Habitat/Hydrology	Dam	Not Described	Quantitative	Random Sample	None	0	0	0	Not Described	Not Described	Mixed-Effects	No
29	Habberfield et al. 2014	Journal of the American Water Resources Association	Rapid Geomorphic and Habitat Stream Assessment Techniques Inform Restoration Differently Based on Levels of Stream Disturbance	United States/New York	Freshwater	Stream	Ecosystem Function	In-Channel Habitat	Multiple Stressors	None	Qualitative	RC	Channel Stability Ranking Scheme/Pfankuch Channel Stability Evaluatoin Proceedure/USEPA Rapid Bioassessment Protocols for Streams and Rivers	0	0	0	2	1	Anova	Yes
30	Harper et al. 1998	Aquatic Conservation: Marine and Freshwater Ecosystems	Artificial riffles in river rehabilitation: Setting the goals and measuring the successes	United Kingdom	Freshwater	Stream	Ecosystem Function	Geomorphology/Habitat Types/Macroinvertebrates	No Development	Artificial Riffles	Quantitative	Targeted	None	0	1	3	0	20	Pearson's Correlation	Yes

Reference ID	Reference	Publisher	Title	Location	Ecosystem	Habitat Type	Habitat Function	Biological Physical Response	Development Type	Intervention	Data Type	Monitoring Design	Standardized Protocol	Before Monitoring Years	After Monitoring Years	Years Post Intervention	Number of Control Sites	Number of Altered Sites	Statistical Analyses	Rapid < 1 day
31	Heady et al. 2015	Ecological Indicators	Assessing California's bar-built estuaries using the California Rapid Assessment Method	United States/California	Freshwater	Estuary	Nursery	Biotic Structure/Buffer and Landscape Context/Hydrology/Physical Habitat	No Development	None	Qualitative	Random Sample	California Rapid Assessment Method	0	0	0	0	32	Spearman's Rank Correlation	Yes
32	Hilderbrand et al. 1997	CJFAS	Effects of large woody debris placement on stream channels and benthic macroinvertebrates	United States/Virginia	Freshwater	Stream	Ecosystem Function	Habitat Types/Macroinvertebrates	Deforestation	Large Wood Addition	Quantitative	CI	None	1	2	0	2	2	Anova	Yes
33	Howell et al. 2012	Restoration Ecology	Responses of Fish to Experimental Introduction of Structural Woody Habitat in Riffles and Pools	Australia	Freshwater	Stream	Ecosystem Function	Fish Assemblage	Multiple Stressors	Large Wood Addition	Quantitative	BACI	None	1	0	1	3	3	Anosim	Unclear
34	Howson et al. 2009	River Research and Applications	Fish assemblage response to rehabilitation of a sand-slugged lowland river	Australia	Freshwater	Stream	Ecosystem Function	Fish Assemblage	Multiple Stressors	Large Wood Addition/Sediment Removal	Quantitative	BACI	None	1	1	2	3	1	Permanova	No
35	Ilmonen et al. 2012	Freshwater Science	Responses of spring macroinvertebrate and bryophyte communities to habitat modification: community composition, species richness, and red-listed species	European Union/Finland	Freshwater	Stream	Ecosystem Function	Bryophyte/Macroinvertebrates	Multiple Stressors	None	Quantitative	RC	None	0	0	0	55	20	Beast	No
36	Johnson and Ringler 2014	Ecological Indicators	The response of fish and macroinvertebrate assemblages to multiple stressors: A comparative analysis of aquatic communities in a perturbed watershed (Onondaga Lake, NY)	United States/New York	Freshwater	Stream	Ecosystem Function	Fish Assemblage/Macroinvertebrates	Multiple Stressors	None	Mix	Random Sample	New York State Department of Environmental Conservation - Macroinvertebrates/USEPA Rapid Bioassessment Protocols for Streams and Rivers	0	0	0	0	17	Pearson's Correlation	No
37	Jun et al. 2012	International Journal of Environmental Research and Public Health	A Multimetric Benthic Macroinvertebrate Index for the Assessment of Stream Biotic Integrity in Korea	Korea	Freshwater	Stream	Ecosystem Function	Macroinvertebrates	No Development	None	Quantitative	RC	Korean Benthic Macroinvertebrate Index of Biological Integrity	0	0	0	112	276	PCA	Yes
38	Kireta et al. 2012	Ecological Indicators	Planktonic and periphytic diatoms as indicators of stress on great rivers of the	United States	Freshwater	River	Ecosystem Function	Periphyton Diatoms/Phytoplankton Diatoms	Multiple Stressors	None	Quantitative	Random Sample	None	0	0	0	0	184	Canocial Correspondence Analysis	Unclear

Reference ID	Reference	Publisher	Title	Location	Ecosystem	Habitat Type	Habitat Function	Biological Physical Response	Development Type	Intervention	Data Type	Monitoring Design	Standardized Protocol	Before Monitoring Years	After Monitoring Years	Years Post Intervention	Number of Control Sites	Number of Altered Sites	Statistical Analyses	Rapid < 1 day
			United States: Testing water quality and disturbance models																	
39	Korte et al. 2010	Hydrobiologia	Assessing river ecological quality using benthic macroinvertebrates in the Hindu Kush-Himalayan region	Bangladesh/Bhutan/India/Nepal/Pakistan	Freshwater	Stream	Ecosystem Function	Macroinvertebrates	Multiple Stressors	None	Quantitative	Random Sample	USEPA Rapid Bioassessment Protocols for Streams and Rivers	0	0	0	0	198	PCA	Yes
40	Krasnicki et al. 2001	Tasmania Department fo Primary Industries, Water and Environment, Hobart	Australia-Wide Assessment of River Health : Tasmanian Bioassessment Report	Tasmania	Freshwater	Stream	Ecosystem Function	Physical Habitat/Macroinvertebrates	Multiple Stressors	None	Mix	RC	Ausriva/USEPA Rapid Bioassessment Protocols for Streams and Rivers	0	0	0	216	257	None	Yes
41	Langer and Smith 2001	Regulated Rivers: Research and Management	Effects of habitat enhancement on 0-group fishes in a lowland river	United Kingdom	Freshwater	Stream	Spawning/Rearing	Fish Productivity	Multiple Stressors	Bank Reconstruction	Quantitative	CI	None	0	0	0	4	4	Anova	No
42	Leps et al. 2016	Science of the Total Environment	Time is no healer: increasing restoration age does not lead to improved benthic invertebrate communities in restored river reaches	European Union/Germany	Freshwater	Stream	Ecosystem Function	Hydromorphology/Macroinvertebrates	Multiple Stressors	Various	Quantitative	CI	AQEM-STAR	0	0	1-25	4	4	Permanova	Yes
43	Lopez and Fennessy 2002	Ecological Applications	Testing the Floristic Quality Assessment Index as an Indicator of Wetland Condition	United States/Ohio	Freshwater	Wetland	Ecosystem Function	Floristic Assemblage	Multiple Stressors	None	Quantitative	Random Sample	FQAI	0	0	0	0	20	Spearman's Rank Correlation	No
44	McCaffery et al. 2007	Transactions of the American Fisheries Society	Effects of Road Decommissioning on Stream Habitat Characteristics in the South Fork Flathead River, Montana	United States/Montana	Freshwater	Stream	Spawning/Rearing	Sediment	Roads	Road Decommission	Quantitative	Random Sample	None	0	0	0	6	6	Kruskal-Wallis Test/Mann-Whitney Test/Pearson's Correlation	Unclear
45	Menetry et al. 2011	Ecological Indicators	The CIEPT: A macroinvertebrate-based multimetric index for assessing the ecological quality of Swiss lowland ponds	European Union/Switzerland	Freshwater	Pond	Ecosystem Function	Amphibian Assemblage/Macroinvertebrates/ Vegetation Index	No Development	Not Described	Quantitative	Random Sample	CIEPT	0	0	0	17	7	Pearson's Correlation	No
46	Miler et al. 2013	Ecological Indicators	Morphological alterations of lake shores in Europe : A multimetric ecological assessment	European Union	Freshwater	Lake	Ecosystem Function	Physical Habitat/Macroinvertebrates	Multiple Stressors	None	Quantitative	RC	Lake Habitat Survey/Littoral Invertebrate Multimetric Index Based on Habitat Sampling/Littoral Invertebrate	0	0	0	0	51	Anosim/Permanova	Yes

Reference ID	Reference	Publisher	Title	Location	Ecosystem	Habitat Type	Habitat Function	Biological Physical Response	Development Type	Intervention	Data Type	Monitoring Design	Standardized Protocol	Before Monitoring Years	After Monitoring Years	Years Post Intervention	Number of Control Sites	Number of Altered Sites	Statistical Analyses	Rapid < 1 day
			approach using benthic macroinvertebrates										Multimetric Index Based on Composite Sampling							
47	Miller et al. 2004	Journal of the American Water Resources Association	STREAM ASSESSMENTS USING BIOTIC INDICES: RESPONSES TO PHYSICOCHEMICAL VARIABLES	United States/Wyoming	Freshwater	Stream	Ecosystem Function	Physical Habitat/Macroinvertebrates/Water Chemistry	Multiple Stressors	Not Described	Quantitative	Random Sample	USEPA Rapid Bioassessment Protocols for Streams and Rivers	0	0	0	0	9	PCA/Pearson's Correlation	Yes
48	Mollard et al. 2013	Wetlands	Monitoring and Assessment of Wetland Condition Using Plant Morphologic and Physiologic Indicators	Canada/Alberta	Freshwater	Wetland	Ecosystem Function	Plant Assemblages	Agriculture/Urbanization	Not Described	Quantitative	Targeted	None	0	0	0	8	25	PCA/Spearman's Rank Correlation	Yes
49	Muehlbauer et al. 2009	Hydrobiologia	Short-term responses of decomposers to flow restoration in Fossil Creek, Arizona, USA	United States/Arizona	Freshwater	Stream	Ecosystem Function	Fungi/Leaf Decomposition/Macroinvertebrates/Water Chemistry	Dam	Flow Restoration	Quantitative	BACI	None	1	1	0.5	1	1	Anova	Unclear
50	Nathan et al. 2018	Ecological Indicators	Are culvert assessment scores an indicator of Brook Trout Salvelinus fontinalis population fragmentation?	United States/Connecticut	Freshwater	Stream	Migration	Fish Passage	Roads	None	Mix	Targeted	North Atlantic Aquatic Connectivity Collaborative Protocol	0	0	0	11	17	Pearson's Correlation/T-Test	Yes
51	Neto et al. 2013	Ecological Indicators	Seagrass Quality Index (SQI), a Water Framework Directive compliant tool for the assessment of transitional and coastal intertidal areas	European Union/Portugal	Marine	Estuary	Ecosystem Function	Sea Grass Assemblage	Multiple Stressors	Flow Restoration	Quantitative	BA	Sea Grass Quality Index	12	12	1-12	0	1	Spearman's Rank Correlation	Unclear
52	Pallottini et al. 2017	Inland Waters	An efficient semi-quantitative macroinvertebrate multimetric index for the assessment of water and sediment contamination in streams	European Union/Italy	Freshwater	Stream	Ecosystem Function	Macroinvertebrates	Multiple Stressors	None	Mix	RC	Semi-Quantitative Multimetric Index	0	0	0	10	11	Kolmogorov-Smirnov Test	Yes
53	Pont et al. 2006	Journal of Applied Ecology	Assessing river biotic condition at a continental scale: a European approach using functional metrics and fish assemblages	European Union	Freshwater	Stream	Ecosystem Function	Fish Assemblage	Multiple Stressors	None	Quantitative	RC	None	0	0	0	1608	5252	Probabilities/Regression	Unclear
54	Pont et al. 2007	Fisheries Managem	Development of a fish-based index for the	European Union	Freshwater	Stream	Ecosystem Function	Fish Assemblage	Multiple Stressors	None	Quantitative	RC	None	0	0	0	1608	5252	Probabilities/Regression	Unclear

Reference ID	Reference	Publisher	Title	Location	Ecosystem	Habitat Type	Habitat Function	Biological Physical Response	Development Type	Intervention	Data Type	Monitoring Design	Standardized Protocol	Before Monitoring Years	After Monitoring Years	Years Post Intervention	Number of Control Sites	Number of Altered Sites	Statistical Analyses	Rapid < 1 day
		ent and Ecology	assessment of river health in Europe: the European Fish Index																	
55	Porst et al. 2012	Fundamental and Applied Limnology	Efficient sampling methodologies for lake littoral invertebrates in compliance with the European Water Framework Directive	European Union	Freshwater	Lake	Ecosystem Function	Macroinvertebrates	Multiple Stressors	None	Quantitative	RC	None	0	0	0	14	19	Anosim/NMDS/Permanova	Yes
56	Radwell and Kwak 2005	Environmental Management	Assessing Ecological Integrity of Ozark Rivers to Determine Suitability for Protective Status	United States/Arkansas	Freshwater	Stream	Ecosystem Function	Physical Habitat/Fish Assemblage/Macroinvertebrates/ Watershed Attributes/Water Chemistry	Multiple Stressors	None	Quantitative	RC	IBI	0	0	0	5	5	Cluster Analysis/Ordinal Scaling	Unclear
57	Raposa et al. 2017	Estuaries and Coasts	Evaluating Tidal Wetland Restoration Performance Using National Estuarine Research Reserve System Reference Sites and the Restoration Performance Index (RPI)	United States	Marine	Wetland	Ecosystem Function	Hydrology/Physical Habitat/Soils/Vegetation	Multiple Stressors	Various	Quantitative	RC	Restoration Performance Index	0	1-4	0-14	9	17	Anosim/Beast/Regression/ RPI/Simpler	Unclear
58	Rehn 2009	River Research and Applications	Benthic macroinvertebrates as indicators of biological condition below hydropower dams on west slope Sierra Nevada streams, California, USA	United States/California	Freshwater	Stream	Ecosystem Function	Physical Habitat/Macroinvertebrates/Periphyton/Water Chemistry	Dam	None	Mix	RC	USEPA Rapid Bioassessment Protocols for Streams and Rivers	0	0	Many	20	54	Regression	Unclear
59	Robertson et al. 2016	Ecological Indicators	Optimising a widely-used coastal health index through quantitative ecological group classifications and associated thresholds	New Zealand	Marine	Estuary	Ecosystem Function	Benthic Condition/Macroinvertebrates	Multiple Stressors	None	Mix	RC	AMBI biotic index	0	0	0	0	21	Regression/Regression Trees	Unclear
60	Rowe et al. 2009	Environmental Management	Evaluating stream restoration: A case study from two partially developed 4th order Connecticut, U.S.A. streams and evaluation monitoring strategies	New Zealand	Freshwater	Stream	Ecosystem Function	Aquatic Biodiversity Intact/Aquatic Fauna Intact/Fish Fauna Intact/Fish Spawning Habitat Intact/Macroinvertebrate Fauna Intact/Riparian Vegetation Intact	Urbanization	Not Described	Mix	RC	Yes	0	1	0	Not Described	21	Not Described	Yes

Reference ID	Reference	Publisher	Title	Location	Ecosystem	Habitat Type	Habitat Function	Biological Physical Response	Development Type	Intervention	Data Type	Monitoring Design	Standardized Protocol	Before Monitoring Years	After Monitoring Years	Years Post Intervention	Number of Control Sites	Number of Altered Sites	Statistical Analyses	Rapid < 1 day
61	Schiff et al. 2011	River Research and Applications	Evaluating stream restoration: A case study from two partially developed 4th order Connecticut, U.S.A. streams and evaluation monitoring strategies	United States	Freshwater	Stream	Ecosystem Function	Physical Habitat/Macroinvertebrates/Water Chemistry	Multiple Stressors	Various	Quantitative	RC	USEPA Rapid Bioassessment Protocols for Streams and Rivers	0	0	Many	13	14	Anova	Unclear
62	Schmutz et al. 2016	Hydrobiologia	Response of fish assemblages to hydromorphological restoration in central and northern European rivers	European Union	Freshwater	Stream	Ecosystem Function	Fish Assemblage/Hydromophology	Multiple Stressors	Instream Structures/Remenandering/Widening	Quantitative	CI	None	0	0	3-20	15	15	Classification Regression Trees/T-Test	Unclear
63	Smith et al. 2007	Biodiversity and Conservation	Assessing Riparian Quality Using Two Complementary Sets of Bioindicators	South Africa	Freshwater	Stream	Ecosystem Function	Macroinvertebrates/Riparian Vegetation	Multiple Stressors	None	Quantitative	Targeted	SASS5 Protocol	0	0	0	0	70	Canocial Correspondence Analysis/Regression	Yes
64	Stander and Ehrenfeld 2009	Wetlands	Rapid assessment of urban wetlands: Functional assessment model development and evaluation	United States/New Jersey	Freshwater	Wetland	Ecosystem Function	Nitrogen Cycling	Multiple Stressors	None	Quantitative	Targeted	Functional Capacity Index	0	0	0	5	9	Regression	Yes
65	Suir and Sasser 2017	US Army Corps of Engineers	Floristic Quality Index of Restored Wetlands in Coastal Louisiana Environmental Laboratory	United States/Louisiana	Freshwater	Wetland	Ecosystem Function	Vegetation	Hurricanes and Salinity Spikes	None	Quantitative	RC	Foristic Quality Index	0	0	0	117	442	None	Yes
66	Talman et al. 1996	Tasmania Department of Enviornment and Land Management	Mount Lyell Remediation: Monitoring of benthic invertebrates in Macquaire Harbour, western Tasmania	Tasmania	Marine	Near-shore	Ecosystem Function	Physical Habitat/Bivalves/Macroinvertebrates	Mine	Improve Water Quality and Sediment Quality	Quantitative	Random Sample	None	0	0	0	0	38	Anova	Yes
67	Tornblom and Anglstam 2011	Fundamental and Applied Limnology	Rapid assessment of headwater stream macroinvertebrate diversity: an evaluation of surrogates across a land-use gradient	European Union/Poland/Romania/Ukraine	Freshwater	Stream	Ecosystem Function	Macroinvertebrates	Multiple Stressors	None	Quantitative	Targeted	Czech Saprobic Index	0	0	0	0	25	Regression	Yes
68	Twohig and Stolt 2011	Wetlands	Soils-based rapid assessment for quantifying changes in salt marsh condition as a result of	United States/Rhode Island/Massachusetts	Marine	Wetland	Ecosystem Function	Soil Properties	Tidal Restriction	None	Quantitative	RC	None	0	0	0	2	2	T-Test	Yes

Reference ID	Reference	Publisher	Title	Location	Ecosystem	Habitat Type	Habitat Function	Biological Physical Response	Development Type	Intervention	Data Type	Monitoring Design	Standardized Protocol	Before Monitoring Years	After Monitoring Years	Years Post Intervention	Number of Control Sites	Number of Altered Sites	Statistical Analyses	Rapid < 1 day
			hydrologic alteration																	
69	Valentine-Rose and Layman 2011	Restoration Ecology	Response of Fish Assemblage Structure and Function Following Restoration of Two Small Bahamian Tidal Creeks	Bahamas	Marine	Estuary	Ecosystem Function	Fish Assemblage	Sedimentation	Hydraulic Connectivity	Quantitative	BACI	None	1	2	0-2	2	2	Anosim/NMDS/Simper	No
70	Vander Laan et al. 2013	Freshwater Science	Linking land use, in-stream stressors, and biological condition to infer causes of regional ecological impairment in streams	United States/Nevada	Freshwater	Stream	Ecosystem Function	Macroinvertebrates	Multiple Stressors	None	Quantitative	RC	None	0	0	0	165	401	Random Forests	Yes
71	Vehanen et al. 2010	Freshwater Biology	Effects of habitat rehabilitation on brown trout (Salmo trutta) in boreal forest streams	European Union/Finland	Freshwater	Stream	Spawning/Rearing	Fish Productivity	Multiple Stressors	Bolder Addition/Large Wood Addition	Quantitative	BACI	None	3	3	0	6	12	Anova	Yes
72	Verissimo et al. 2012	Ecological Indicators	Ability of benthic indicators to assess ecological quality in estuaries following management	European Union/Portugal	Marine	Estuary	Ecosystem Function	Macroinvertebrates	Multiple Stressors	Hydraulic Connectivity	Quantitative	BA	None	3	2	0	Not Described	15	Bray-Curtis/NMDS/PCA/Permanova	Unclear
73	Walker and MacAskill 2014	Environmental Monitoring and Assessment	Monitoring water quality in Sydney Harbour using blue mussels during remediation of the Sydney Tar Ponds, Nova Scotia, Canada	Canada/Nova Scotia	Marine	Near-shore	Ecosystem Function	Crab Tissue/Mussels/Sediment/Water Quality	Contaminants Site Remediation	Not Described	Quantitative	BA	None	1	3	0	0	11	Anova	No
74	Weigel and Dimick 2011	Journal of the North American Bethological Society	Development, validation, and application of a macroinvertebrate-based Index of Biotic Integrity for nonwadeable rivers of Wisconsin	United States/Wisconson	Freshwater	River	Ecosystem Function	Fish Assemblage/Macroinvertebrates	Multiple Stressors	None	Quantitative	RC	None	0	0	0	32	68	Anova	No
75	Wellnitz et al. 2014	Limnologica	Do installed stream logjams change benthic community structure?	United States/Minnesota	Freshwater	Stream	Ecosystem Function	Macroinvertebrates	Unknown	Large Wood Addition	Quantitative	BACI	None	1	1	0	2	2	Manova	Yes
76	Wigand et al. 2011	Environmental Monitoring and Assessment	Development and validation of rapid assessment indices of condition for coastal tidal wetlands in southern New England, USA	United States/Connecticut/Massachusetts/Rhode Island	Marine	Wetland	Ecosystem Function	Disturbance/Soil/Vegetation	Multiple Stressors	None	Quantitative	RC	None	0	0	0	10	71	Regression/PCA	Yes

Reference ID	Reference	Publisher	Title	Location	Ecosystem	Habitat Type	Habitat Function	Biological Physical Response	Development Type	Intervention	Data Type	Monitoring Design	Standardized Protocol	Before Monitoring Years	After Monitoring Years	Years Post Intervention	Number of Control Sites	Number of Altered Sites	Statistical Analyses	Rapid < 1 day
77	Winger et al. 2005	Environmental Monitoring and Assessment	Combined use of rapid bioassessment protocols and sediment quality triad to assess stream quality	United States	Freshwater	Stream	Ecosystem Function	Physical Habitat/Macroinvertebrates/Fish/Sediment	Waste Water Treatment Facility	None	Quantitative	RC	Sediment Quality Triad/USEPA Rapid Bioassessment Protocols for Streams and Rivers	0	0	0	1	4	Anova	Yes

Table A 3. List of literature reviewed for research document.

Reference ID	Full Citation
1	Adomato, T., Fish, U.S., Service, W., Ecological, O., and Field, S. 1997. The use of rapid bioassessment protocols to describe fish and benthic macroinvertebrate communities in three creeks near the Little River National Wildlife Refuge, Mccurtain County Oklahoma. (June).
2	Albertson, L.K., Koenig, L.E., Lewis, B.L., Zeug, S.C., Harrison, L.R., and Cardinale, B.J. 2013. How does restored habitat for chinook salmon (oncorhynchus tshawytscha) in the merced river in california compare with other chinook streams? River Res. Appl. 29(4): 469–482. doi:10.1002/rra.1604.
3	Alford, J.B. 2014. Multi-scale assessment of habitats and stressors influencing stream fish assemblages in the Lake Pontchartrain Basin, USA. Hydrobiologia 738(1): 129–146. doi:10.1007/s10750-014-1925-2.
4	Angradi, T.R., Bolgrien, D.W., Jicha, T.M., Pearson, M.S., Hill, B.H., Taylor, D.L., Schweiger, E.W., Shepard, L., Batterman, A.R., Moffett, M.F., Elonen, C.M., and Anderson, L.E. 2009. A bioassessment appro
5	Antón, A., Elosegi, A., García-Arberas, L., Díez, J., and Rallo, A. 2011. Restoration of dead wood in Basque stream channels: Effects on brown trout population. Ecol. Freshw. Fish 20(3): 461–471. WILEY-BLACKWELL, 111 RIVER ST, HOBOKEN 07030-5774, NJ USA. doi:10.1111/j.1600-0633.2010.00482.x.
6	Arthur, A., and Kauss, P. 2000. Sediment and benthic community assessment of the St. Marys River.
7	Benedetti-Cecchi, L., and Osio, G.C. 2007. Replication and mitigation of effects of confounding variables in environmental impact assessment: Effect of marinas on rocky-shore assemblages. Mar. Ecol. Prog. Se
8	Bennett, S.A. 2009. Bioassessment of streams in the northwest BC using the Skeena Beast09. Gitxsan For. Enterp. Westfraser Mills Ltd. (March): 1–76.
9	Berkowitz, J.F., and White, J.R. 2013. Linking wetland functional rapid assessment models with quantitative hydrological and biogeochemical measurements across a restoration chronosequence. Soil Sci. Soc. Am
10	Bonada, N., Dallas, H., Rieradevall, M., Prat, N., and Day, J. 2006. A comparison of rapid bioassessment protocols used in 2 regions with Mediterranean climates, the Iberian Peninsula and South Africa. J. North Am. Benthol. Soc. 25(2): 487–500.
11	Borisko, J.P., Kilgour, B.W., Stanfield, L.W., and Jones, F.C. 2007. An evaluation of rapid bioassessment protocols for stream benthic invertebrates in Southern Ontario, Canada. Water Qual. Res. J. Canada 42(3): 184–193.
12	Boys, C.A., Kroon, F.J., Glasby, T.M., and Wilkinson, K. 2012. Improved fish and crustacean passage in tidal creeks following floodgate remediation. J. Appl. Ecol. 49(1): 223–233. doi:10.1111/j.1365-2664.2011.02101.x.
13	Brown, L.E., Ramchunder, S.J., Beadle, J.M., and Holden, J. 2016. Macroinvertebrate community assembly in pools created during peatland restoration. Sci. Total Environ. 569–570: 361–372. doi:10
14	Cahill, C.L., Erwin, A.C., Howland, K.L., Hulsman, M.F., Lunn, B.D., Noddin, F., Tonn, W.M., Baki, A.B., Courtice, G., and Zhu, D.Z. 2015. Assessing responses of fish to habitat enhancement in Barrenlands streams of the Northwest Territories. North Am. J. Fish. Manag. 35(4): 755–764. doi:10.1080/02755947.2015.1044626.
15	Chamberlain, S.J., and Brooks, R.P. 2016. Testing a rapid floristic quality index on headwater wetlands in Central Pennsylvania, USA. Ecol. Indic. 60: 1142–1149. doi:10.1016/j.ecolind.2015.09.004.

Reference ID	Full Citation
16	Cianfrani, C., Hession, W.C., McBride, M., and Pizzuto, J.E. 2001. Assessment of urban streams in Fairmount Park, Philadelphia, PA. In Bridging the Gap. American Society of Civil Engineers. pp. 1–7. doi:10.1061/40569(2001)117.
17	Cohen, M.J., Lane, C.R., Reiss, K.C., Surdick, J.A., Bardi, E., and Brown, M.T. 2005. Vegetation based classification trees for rapid assessment of isolated wetland condition. Ecol. Indic. 5(3): 189–206. doi:10.1016/j.ecolind.2005.01.002.
18	Cooperman, M.S., Hinch, S.G., Bennett, S., Branton, M.A., Galbraith, R. V, Quigley, J.T., and Heise, B.A. 2007. Streambank restoration effectiveness : lessons learned from a comparative study. Fish. Mag. 32(6): 278–291.
19	Cordell, J.R., and Toft, J. 2012. 2010 Invertebrate monitoring at Duwamish Waterway restoration sites : Hamm Creek, Herring's House, Northwind's Weir, and Kenco Marine. United States Fish Wildl. Serv.
20	de Bikuña, B.G., López, E., Leonardo, J.M., Arrate, J., Martínez, A., and Manzanos, A. 2015. Development of a multimetric benthic macroinvertebrate index for assessing the ecological condition of Basque streams (north of Spain). Fundam. Appl. Limnol. / Arch. für Hydrobiol. 187(1): 21–32. doi:10.1127/fal/2015/0741.
21	de Mutsert, K., and Cowan, J.H. 2012. A before–after–control–impact analysis of the effects of a Mississippi River freshwater diversion on estuarine nekton in Louisiana, USA. Estuaries and Coasts 35(5
22	Doll, B.A., Jennings, G.D., Spooner, J., Penrose, D.L., and Usset, J.L. 2015. Evaluating the eco-geomorphological condition of restored streams using visual assessment and macroinvertebrate metrics. J. Am. Water Resour. Assoc. 51(1): 68–83. doi:10.1111/jawr.12233.
23	Doll, B., Jennings, G., Spooner, J., Penrose, D., Usset, J., Blackwell, J., and Fernandez, M. 2016. Can rapid assessments predict the biotic condition of restored streams? Water (Switzerland) 8(4): 143. doi:10.3390/w8040143.
24	Doll, B., Jennings, G., Spooner, J., Penrose, D., Usset, J., Blackwell, J., and Fernandez, M. 2016. Identifying watershed, landscape, and engineering design factors that influence the biotic condition of restored streams. Water (Switzerland) 8(4): 1–18. doi:10.3390/w8040151.
25	Entrekin, S.A., Tank, J.L., Rosi-Marshall, E.J., Hoellein, T.J., and Lamberti, G.A. 2009. Response of secondary production by macroinvertebrates to large wood addition in three Michigan streams. Freshw. Biol. 54(8): 1741–1758. doi:10.1111/j.1365-2427.2009.02223.x.
26	Flores, L., Giorgi, A., González, J.M., Larrañaga, A., Díez, J.R., and Elosegi, A. 2017. Effects of wood addition on stream benthic invertebrates differed among seasons at both habitat and reach scales. Ecol. Eng. 106(A): 116–123. doi:10.1016/j.ecoleng.2017.05.036.
27	Franklin, P.A., and Bartels, B. 2012. Restoring connectivity for migratory native fish in a New Zealand stream: effectiveness of retrofitting a pipe culvert. Aquat. Conserv. Mar. Freshw. Ecosyst. 22(4): 489–497. doi:10.1002/aqc.2232.
28	Goodman, D.H., Som, N.A., Alvarez, J., and Martin, A. 2015. A mapping technique to evaluate age-0 salmon habitat response from restoration. Restor. Ecol. 23(2): 179–185. doi:10.1111/rec.12148.
29	Habberfield, M.W., Blersch, S.S., Bennett, S.J., and Atkinson, J.F. 2014. Rapid geomorphic and habitat stream assessment techniques inform restoration differently based on levels of stream disturbance. J. Am. Water Resour. Assoc. 50(4): 1051–1062. doi:10.1111/jawr.12156.
30	Harper, D., Ebrahimnezhad, M., and Climent I Cot, F. 1998. Artificial riffles in river rehabilitation: setting the goals and measuring the successes. Aquat. Conserv. Mar. Freshw. Ecosyst. 8(1): 5–16. doi:10.1002/(SICI)1099-0755(199801/02)8:1<5::AID-AQC266>3.0.CO;2-G.

Reference ID	Full Citation
31	Heady, W.N., Clark, R.P., O'Connor, K., Clark, C., Endris, C., Ryan, S., and Stoner-Duncan, S. 2015. Assessing California's bar-built estuaries using the California rapid assessment method. Ecol. Indic. 58: 300–310. doi:10.1016/j.ecolind.2015.05.062.
32	Hilderbrand, R.H., Lemly, A.D., Dolloff, C.A., and Harpster, K.L. 1997. Effects of large woody debris placement on stream channels and benthic macroinvertebrates. Can. J. Fish. Aquat. Sci. 54(4): 931–939. doi:10.1139/f96-334.
33	Howell, T.D., Arthington, A.H., Pusey, B.J., Brooks, A.P., Creese, B., and Chaseling, J. 2012. Responses of fish to experimental introduction of structural woody habitat in riffles and pools. Restor. Ecol. 20(1): 43–55. doi:10.1111/j.1526-100X.2010.00747.x.
34	Howson, T.J., Robson, B.J., and Mitchell, B.D. 2009. Fish assemblage response to rehabilitation of a sand-slugged lowland river. River Res. Appl. 25(10): 1251–1267. doi:10.1002/rra.1226.
35	Ilmonen, J., Mykrä, H., Virtanen, R., Paasivirta, L., and Muotka, T. 2012. Responses of spring macroinvertebrate and bryophyte communities to habitat modification: community composition, species richness, and red-listed species. Freshw. Sci. 31(2): 657–667. doi:10.1899/10-060.1.
36	Johnson, S.L., and Ringler, N.H. 2014. The response of fish and macroinvertebrate assemblages to multiple stressors: a comparative analysis of aquatic communities in a perturbed watershed (Onondaga Lake, NY). Ecol. Indic. 41: 198–208. doi:10.1016/j.ecolind.2014.02.006.
37	Jun, Y.-C., Won, D.-H., Lee, S.-H., Kong, D.-S., and Hwang, S.-J. 2012. A Multimetric Benthic Macroinvertebrate Index for the Assessment of Stream Biotic Integrity in Korea. Int. J. Environ. Res. Public Health 9(12): 3599–3628. MDPI AG, POSTFACH, CH-4005 BASEL, SWITZERLAND. doi:10.3390/ijerph9103599.
38	Kireta, A.R., Reavie, E.D., Sgro, G. V., Angradi, T.R., Bolgrien, D.W., Hill, B.H., and Jicha, T.M. 2012. Planktonic and periphytic diatoms as indicators of stress on great rivers of the United States: testing water quality and disturbance models. Ecol. Indic. 13(1): 222–231. doi:10.1016/j.ecolind.2011.06.006.
39	Korte, T., Baki, A.B.M., Ofenböck, T., Moog, O., Sharma, S., and Hering, D. 2010. Assessing river ecological quality using benthic macroinvertebrates in the Hindu Kush-Himalayan region. Hydrobiologia 651(1): 59–76. doi:10.1007/s10750-010-0290-z.
40	Krasnicki, A.T., Pinto, R., Read, M., and Krasnicki, T. 2001. Australia-Wide assessment of river health : Tasmanian bioassessment report (TAS final report). Environ. Aust. (5): 65.
41	Langler, G.J., and Smith, C. 2001. Effects of habitat enhancement on 0-group fishes in a lowland river. Regul. Rivers Res. Manag. 17(July 1999): 677–686. doi:10.1002/rrr.627.
42	Leps, M., Sundermann, A., Tonkin, J.D., Lorenz, A.W., and Haase, P. 2016. Time is no healer: increasing restoration age does not lead to improved benthic invertebrate communities in restored river reaches. Sci. Total Environ. 557–558: 722–732. doi:10.1016/j.scitotenv.2016.03.120.
43	Lopez, R.D., and Fennessy, M.S. 2002. Testing the floristic quality assessment index as an indicator of wetland condition. Ecol. Appl. 12(2): 487. doi:10.2307/3060957.
44	McCaffery, M., Switalski, T.A., and Eby, L. 2007. Effects of road decommissioning on stream habitat characteristics in the South Fork Flathead River, Montana. Trans. Am. Fish. Soc. 136(3): 553–561. doi:10.1577/T06-134.1.

Reference ID	Full Citation
45	Menetrey, N., Oertli, B., and Lachavanne, J.B. 2011. The CIEPT: A macroinvertebrate-based multimetric index for assessing the ecological quality of Swiss lowland ponds. Ecol. Indic. 11(2): 590–600. doi:10.1016/j.ecolind.2010.08.005.
46	Miler, O., Porst, G., Mcgoff, E., Pilotto, F., Donohue, L., Jurca, T., Solimini, A., Sandin, L., Irvine, K., Aroviita, J., Clarke, R., and Pusch, M.T. 2013. Morphological alterations of lake shores in Europe: a multimetric ecological assessment approach using benthic macroinvertebrates. Ecol. Indic. 34: 398–410. doi:https://doi.org/10.1016/j.ecolind.2013.06.002.
47	Miller, S.W., Skinner, Q.D., and Ready, K.J. 2004. Stream assessments using biotic indices: responses to physicochemical variables. J. Am. Water Resour. Assoc. 40(5): 1173–1188. doi:10.1111/j.1752-1688.2004.tb01577.x.
48	Mollard, F.P.O., Foote, A.L., Wilson, M.J., Crisfield, V., and Bayley, S.E. 2013. Monitoring and assessment of wetland condition using plant morphologic and physiologic indicators. Wetlands 33(5): 939
49	Muehlbauer, J.D., LeRoy, C.J., Lovett, J.M., Flaccus, K.K., Vlieg, J.K., and Marks, J.C. 2009. Short-term responses of decomposers to flow restoration in Fossil Creek, Arizona, USA. Hydrobiologia 618(
50	Nathan, L.R., Smith, A.A., Welsh, A.B., and Vokoun, J.C. 2018. Are culvert assessment scores an indicator of Brook Trout Salvelinus fontinalis population fragmentation? Ecol. Indic. 84: 208–217. doi:1
51	Neto, J.M., Barroso, D. V., and Barría, P. 2013. Seagrass quality index (SQI), a water framework directive compliant tool for the assessment of transitional and coastal intertidal areas. Ecol. Indic. 30</
52	Pallottini, M., Goretti, E., Selvaggi, R., Cappelletti, D., Dedieu, N., and Céréghino, R. 2017. An efficient semi-quantitative macroinvertebrate multimetric index for the assessment of water and sediment con
53	Pont, D., Hugueny, B., Beier, U., Goffaux, D., Melcher, A., Noble, R., Rogers, C., Roset, N., and Schmutz, S. 2006. Assessing river biotic condition at a continental scale: a european approach using function
54	Pont, D., Hugueny, B., and Rogers, C. 2007. Development of a fish-based index for the assessment of river health in Europe: the european fish index. Fish. Manag. Ecol. 14(6): 427–439. doi:10.1111/j.13
55	Porst, G., Miler, O., Donohue, L., Jurca, T., Pilotto, F., Brauns, M., Solimini, A., and Pusch, M. 2016. Efficient sampling methodologies for lake littoral invertebrates in compliance with the european water
56	Radwell, A.J., and Kwak, T.J. 2005. Assessing ecological integrity of Ozark Rivers to determine suitability for protective status. Environ. Manage. 35(6): 799–810. doi:10.1007/s00267-004-0136-4.
57	Raposa, K.B., Lerberg, S., Cornu, C., Fear, J., Garfield, N., Peter, C., Weber, R.L.J., Moore, G., Burdick, D., and Dionne, M. 2017. Evaluating tidal wetland restoration performance using national estuarine
58	Rehn, A.C. 2009. Benthic macroinvertebrates as indicators of biological condition below hydropower dams on west slope Sierra Nevada streams, California, USA. River Res. Appl. 25(2): 208–228. doi:10.10
59	Robertson, B.P., Savage, C., Gardner, J.P.A., Robertson, B.M., and Stevens, L.M. 2016. Optimising a widely-used coastal health index through quantitative ecological group classifications and associated thres
60	Rowe, D.K., Parkyn, S., Quinn, J., Collier, K., Hatton, C., Joy, M.K., Maxted, J., and Moore, S. 2009. A rapid method to score stream reaches based on the overall performance of their main ecological functions. Environ. Manage. 43(6): 1287–1300. SPRINGER, 233 SPRING ST, NEW YORK, NY 10013 USA. doi:10.1007/s00267-009-9302-z.

Reference ID	Full Citation
61	Schiff, R., Benoit, G., and MacBroom, J. 2011. Evaluating stream restoration: A case study from two partially developed 4th order Connecticut, U.S.A. streams and evaluation monitoring strategies. River Res.
62	Schmutz, S., Jurajda, P., Kaufmann, S., Lorenz, A.W., Muhar, S., Paillex, A., Poppe, M., and Wolter, C. 2016. Response of fish assemblages to hydromorphological restoration in central and northern European r
63	Smith, J., Samways, M.J., and Taylor, S. 2007. Assessing riparian quality using two complementary sets of bioindicators. Biodivers. Conserv. 16(9): 2695–2713. doi:10.1007/s10531-006-9081-2.
64	Stander, E.K., and Ehrenfeld, J.G. 2009. Rapid assessment of urban wetlands: Functional assessment model development and evaluation. Wetlands 29(1): 261–276. doi:10.1672/07-145.1.
65	Suir, G.M., and Sasser, C.E. 2017. Floristic quality index of restored wetlands in coastal Louisiana. Ecosyst. Manag. Restor. Res. Progr. Florist. (August): 46.
66	Talman, S., O'Connor, N., Zampatti, B., and Cannan, F. 1996. Mount Lyell remediation: monitoring of benthic invertebrates in Macquaire Harbour, western Tasmania. In Department of Environment and Land
67	Törnblom, J., Roberge, J.-M., and Angelstam, P. 2011. Rapid assessment of headwater stream macroinvertebrate diversity: an evaluation of surrogates across a land-use gradient. Fundam. Appl. Limnol. / Arch. f
68	Twohig, T.M., and Stolt, M.H. 2011. Soils-based rapid assessment for quantifying changes in salt marsh condition as a result of hydrologic alteration. Wetlands 31(5): 955–963. doi:10.1007/s13157-011-0
69	Valentine-Rose, L., and Layman, C.A. 2011. Response of fish assemblage structure and function following restoration of two small Bahamian tidal creeks. Restor. Ecol. 19(2): 205–215. doi:10.1111/j.1526
70	Vander Laan, J.J., Hawkins, C.P., Olson, J.R., and Hill, R.A. 2013. Linking land use, in-stream stressors, and biological condition to infer causes of regional ecological impairment in streams. Freshw. Sci.
71	Vehanen, T., Huusko, A., Mäki-Petäys, A., Louhi, P., Mykrä, H., and Muotka, T. 2010. Effects of habitat rehabilitation on brown trout (Salmo trutta) in boreal forest streams. Freshw. Biol. 55(10): 220
72	Veríssimo, H., Neto, J.M., Teixeira, H., Franco, J.N., Fath, B.D., Marques, J.C., and Patrício, J. 2012. Ability of benthic indicators to assess ecological quality in estuaries following management. Ecol. In
73	Walker, T.R., and MacAskill, D. 2014. Monitoring water quality in Sydney Harbour using blue mussels during remediation of the Sydney Tar Ponds, Nova Scotia, Canada. Environ. Monit. Assess. 186(3): 162
74	Weigel, B.M., and Dimick, J.J. 2011. Development, validation, and application of a macroinvertebrate-based index of biotic integrity for nonwadeable rivers of Wisconsin. J. North Am. Benthol. Soc. 30(
75	Wellnitz, T., Kim, S.Y., and Merten, E. 2014. Do installed stream logjams change benthic community structure? Limnol. - Ecol. Manag. Inl. Waters 49: 68–72. doi:10.1016/j.limno.2014.09.002.
76	Wigand, C., Carlisle, B., Smith, J., Carullo, M., Fillis, D., Charpentier, M., McKinney, R., Johnson, R., and Heltshe, J. 2011. Development and validation of rapid assessment indices of condition for coastal
77	Winger, P. V., Lasier, P.J., and Bogenrieder, K.J. 2005. Combined use of rapid bioassessment protocols and sediment quality triad to assess stream quality. Environ. Monit. Assess. 100(1–3): 267–295. d

APPENDIX B

Inclusion Criteria

1. Is the population or habitat aquatic? ***Yes/No***
2. Is the habitat used by fish spawning, rearing, foraging, or migrating? ***Yes/No***
3. Does the study evaluate mitigation/restoration/offsetting/environmental stressors on habitat function? This includes indicators, indices, or raw data that are used to characterize how habitats function as spawning ground, nursery, rearing, food supply, or migration areas; including if these functions have changed due to some human-induced change to the habitat. ***Yes/No***
4. Does the study use rapid assessment techniques? The study should not rely on extensive data collection (e.g., many years) within a monitoring event. ***Yes/No***
5. Need to answer yes to all for a study to be included. Reviews or assessments of multiple standardized monitoring programs can be put into a Review folder that can be used as needed.
6. Is the study a review or does it demonstrate the development of new methods or metrics? ***Yes/No***

APPENDIX C

Table C 1. Url for government websites and databases. We used a reduced search term list for government databases because most websites did not allow multi-faceted search strings. The reduced search term list included: Aquatic Rapid Assessment Method, Marine "Rapid Assessment" Monitoring, Rapid Bioassessment Protocol, Near Shore Monitoring, Marine "Habitat Assessment", Rapid Assessment Protocol, Estuary Monitoring, and Rapid Lake Assessment.

Source	URL
Web of Science	https://apps.webofknowledge.com
Scopus	https://www.scopus.com/
Canadian Science Library	https://science-libraries.canada.ca/eng/home/
Science.gov	https://www.science.gov/
Australian Government – Department of the Environment and Energy	http://www.environment.gov.au/
New Zealand - Ministry for the Environment	http://www.mfe.govt.nz/
National Oceanic and Atmospheric Administration	http://www.noaa.gov/
European Union – Openaire	https://www.openaire.eu/search/find?keyword
Gov.uk	https://www.gov.uk/
British Columbia – EcoCat	http://a100.gov.bc.ca/pub/acat/public/welcome.do
Alberta Government	https://open.alberta.ca/publications
Ontario Government	http://govdocs.ourontario.ca/

APPENDIX D

We present a table that includes examples of indicators that could be used for functional monitoring. This is not meant to be an exhaustive list, but is to illustrate the types of indictors relevant to functional monitoring of the five habitat types outlined in the *Fisheries Act*. Indicators are broadly categorized into: 1) Physical, 2) Water quantity, 3) Water chemistry, and 4) Biological indicators. Commonly used metrics are also presented. All indictors and metrics can be used across system types unless indicated. Most indicators are based on Smokorowski et al. (2015) with other indicators added based on our review. The amount of variability and spatial (site, reach, watershed) and/or temporal scale (daily, weekly, season) at which the variability occurs will differ among indicators. The use of indicators with high variability should be avoided in a functional monitoring program, as they will require intensive sampling to generate precise estimates. Indicator variability should be considered when selecting indicators (see Indicator Selection section).

Table D 1. Examples of indicators and metrics that could be used for functional monitoring of the five habitats outlined in the Fisheries Act, adapted from (Smokorowski et al. 2015 and references therein). Xs indicate if the indicator (rows) may potentially be useful for assessment of a given habitat (columns). Spawning habitat is defined as habitat that provides the environmental conditions necessary for successful spawning. Nursery habitat is defined as habitat that provides better than average conditions for larval and young-of-the-year fish to grow and survive to the next life stage. Rearing habitat is defined as the area, and the environmental conditions within the area, that support fish growth, survival, and production during the life history stages, from the end of the young-of-the-year or post-larval stage to the adult stage (i.e., usually "juvenile" life stages). Food supply is defined as ecosystem components that contribute to the production of food for fish and are considered fish habitat by the Fisheries Act. Migration habitat is defined as habitat that provides connectivity between all essential habitats (e.g., spawning, nursery, rearing, and food supply habitats), and is required to ensure fish reproduction, growth, and survival at different life stages. This is not an exhaustive list of indicators.

	Indicator	**Metric**	**Spawning**	**Nursery**	**Rearing**	**Food Supply**	**Migration**
Physical	Substrate	% spawning substrate, substrate embeddedness, substrate composition, geometric mean substrate size	X				
	Cover	All systems - % macrophyte cover, bathymetric roughness, density of large woody debris; streams and rivers - % undercut bank, % boulder; marine coastal and estuaries - % macroalgae cover, % biogenic habitats	X	X	X	X	

	Indicator	Metric	Spawning	Nursery	Rearing	Food Supply	Migration
	Habitat type	Area of habitat types: streams - riffles, pools; rivers - slow and fast water; large rivers and lakes - ratio of benthic and pelagic areas; marine coastal and estuary - tide channel, seagrass and marsh habitat		X	X	X	
Water quantity	Water depth	Stream, rivers, lakes - mean depth,% change in water depth between spawning and incubation (may require the use of data loggers)	X	X	X		X
	Water velocity	Streams and rivers - mean, maximum, minimum, gradient	X	X	X		X
	Discharge	Streams and rivers - mean, maximum, minimum (may require the use of data loggers)	X	X	X		X
Water chemistry	Oxygen	Incubation environment dissolved oxygen concentration	X				
		Water column dissolved oxygen concentration		X	X		
	Sediment concentration	Suspended sediment concentration	X	X	X		
		Turbidity	X	X	X		
	Temperature	Mean, maximum, and minimum spawning temperature (may require the use of data loggers)	X	X			X
		Mean, maximum, minimum temperature (may require the use of data loggers)		X	X		

	Indicator	Metric	Spawning	Nursery	Rearing	Food Supply	Migration
	pH	Mean, maximum, minimum		X	X		
	Salinity	Parts per thousand		X	X		
	Nutrients	Nitrogen concentration		X	X	X	
		Phosphorous concentration		X	X	X	
		Periphyton abundance		X	X	X	
Biological	Fish assemblage	Presence of spawning fish	X				X
		Presence of larval or juvenile fish	X	X	X		X

APPENDIX E

We outline a potential approach for the development of a checklist of standardized indicators for the FPP's in-water activities and Pathways of Effects (POE) (DFO 2018b). The checklist consists of a series of potential indicators that could be used to assess changes in habitat associated with each of the endpoints from the Pathways of Effects models (Table E1 and Figure E1). The checklist could be used to assemble a project-specific monitoring protocol. First, the project context is considered, which could include the ecosystem (e.g., freshwater or marine) and types of in-water activities that will be conducted during the project. Based on the in-water activities, the relevant pathways of effects and their endpoints are identified. The user would then check off the standardized indicators that correspond to the identified endpoints and add them to their data collection forms. This approach would not preclude other indicators from being measured, such as the amount of habitat affected by the project and/or indicators specific to species at risk.

For this example, each indicator has been identified for measurement in either freshwater (F), or marine (M), or both (F, M) ecosystems. Indicators could be further categorized by system type (e.g., stream, river, lake, estuary, or marine coastal) and/or habitat function (e.g., spawning, rearing, nursery, food supply, and/or migration habitat). This set of indicators is not complete, and is only for illustrative purposes. Note, 1) additional pathways suggested for some activities (e.g., Dredging – see: fish passage pathway if relevant to changes in hydraulics) have not been considered here, but could be incorporated by working through the Fish Passage POEs; 2) endpoints that examine direct or potential mortality (e.g., In-water activity: Use of industrial equipment, POE endpoint: Potential mortality of fish/egg/ova from equipment), sublethal effects (e.g., In-water activity: Use of explosives, POE endpoint: Lethal or sublethal effects on fish), or other pathway-specific indictors that are not related to habitat function have not been presented but could be considered in future versions of this approach.

Table E 1. Checklist of standardized indicators for in-water activities and Pathways of Effects endpoints (DFO 2018b) that could be used for functional monitoring X's indicate if the indicator (rows) may potentially be useful for assessing a given in-water activity potential effect endpoint (columns).Project POEs would be identified, and for each POE endpoint ecosystem-specific suites of indicators would be assigned to a monitoring protocol. Indicators that are used for marine assessments are denoted by M, and those that are used for freshwater assessments are denoted by F.

Indicator	**Ecosystem**	**∆ in Food Supply**	**∆ in Habitat Structure and Cover**	**∆ in Sediment Concentrations**	**∆ in Nutrient Concentrations**	**∆ in Contaminant Concentrations**	**∆ in Access to Habitats**	**∆ in Salinity**	**∆ in Total Gas Pressure**	**∆ in Thermal Cues or Temperature Barriers**	**∆ in Water Temp**	**∆ in Dissolved O_2**
Nutrients	M,F	X			X							
Substrate	M,F	X	X	X								
Water depth	F		X				X				X	
Water velocity	F		X				X				X	
Fish assemblage	F						X					
Cover	M,F		X								X	
Sediment concentration	M,F			X							X	
Contaminant concentration	M,F					X						
Gradient	F						X					
Fish assemblage	F						X					
Salinity	M,F							X				

Indicator	Ecosystem	∆ in Food Supply	∆ in Habitat Structure and Cover	∆ in Sediment Concentrations	∆ in Nutrient Concentrations	∆ in Contaminant Concentrations	∆ in Access to Habitats	∆ in Salinity	∆ in Total Gas Pressure	∆ in Thermal Cues or Temperature Barriers	∆ in Water Temp	∆ in Dissolved O_2
Dissolved gas pressure	M,F								X			
Temperature	M,F									X	X	
Dissolved O_2 concentration	M,F								X			X

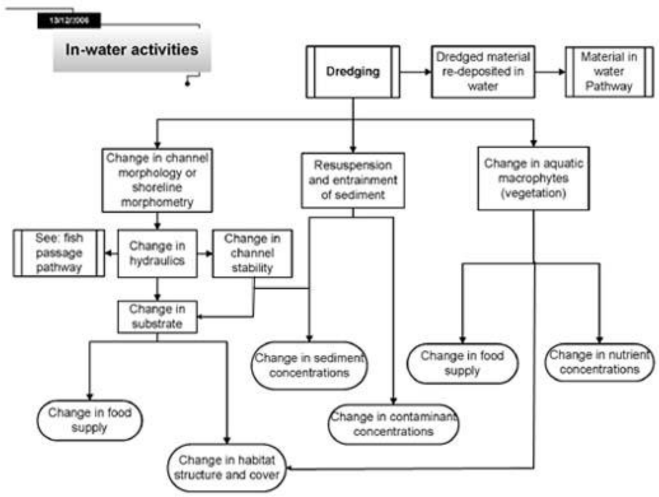


Figure E 1. Pathways of Effects for in-water dredging activities (DFO 2018b). The pathway endpoints (rounded boxes) indicate the final effect of the activity on fish habitat. Indicators are assigned to endpoints in the checklist (Table E1).