3 June 2005

Review of experimental design principles for projects to restore salmon populations in the Arctic-Yukon-Kuskokwim region of Alaska

Randall M. Peterman[1], Brigitte Dorner, Marc A. Nelitz[2]

School of Resource and Environmental Management
Simon Fraser University
Burnaby, British Columbia, Canada, V5A 1S6

[1] Author for correspondence: peterman@sfu.ca

[2] Current address:
ESSA Technologies Ltd.
Suite 300, 1765 West 8th Avenue
Vancouver, B.C., Canada V6J 5C6

A contract report for the
Arctic-Yukon-Kuskokwim Sustainable Salmon Initiative
Anchorage, Alaska

2004-2005 Project # 45237

Please do not cite without the permission of the authors

Table of Contents

Executive Summary / Abstract

This document was written to support the development of a successful restoration program for low-abundance Pacific salmon stocks (*Oncorhynchus* spp.) in the Arctic-Yukon-Kuskokwim (AYK) region of Alaska. By "restoration program" we mean a broad set of human actions, such as new harvest regulations or improvements to salmon habitats, that aim to restore salmon populations to higher abundances or to maintain resilience to future disturbances such as climatic changes. Based on the wide experience with attempts to rebuild salmon stocks throughout the North Pacific Rim over the last 30 years, it is likely that only some of these current and future actions will succeed at their restoration objectives. The reason for doubt is that such actions are vulnerable to unexpected outcomes that result from uncontrollable natural events, as well as our incomplete knowledge of how complex salmon ecosystems function. Actions such as changing mesh size or other harvest regulations, supplementation of juvenile abundance through additions of juvenile fish from incubation boxes or hatcheries, or removal of beaver dams, may have unintended negative effects on target salmon populations through selection for earlier age at maturity, spread of diseases, or loss of deep over-wintering pools for juvenile coho salmon, respectively. Given this uncertainty about outcomes, we should design restoration projects so that their effectiveness can be rigorously evaluated in the future. For example, 5 or 10 years from now we want to be able to convincingly conclude that restoration methods 1, 4, and 5 (whatever those might be) met the restoration objectives, but methods 2 and 3 did not. Future funds could then be focused on methods 1, 4, and 5, unless methods 2 or 3 had offsetting benefits such as fostering of stewardship through local involvement, or improving other ecosystem benefits such as water quality or habitat for other species.

A huge challenge for evaluation of restoration actions is created by the presence of natural variability in ecological systems and simultaneously changing (and potentially confounding) factors. These factors may obscure the actual effects of restoration on salmon populations. To reduce the chance of such confounding, a well-designed action plan could compare, for example, salmon populations from several sites to which a particular type of "treatment" had been applied (e.g. incubation boxes to increase the number of fry produced per egg) with other "control" populations that did not receive that "treatment". Without such comparison groups, incorrect conclusions could easily be drawn about the effectiveness of some restoration action. For example, an increase in abundance of adult salmon recruits that was actually due to improved ocean conditions might be incorrectly attributed to the incubation boxes, or conversely, a lack of increase in abundance caused by deteriorating marine survival rate might lead to an incorrect conclusion that the boxes were not working, when in fact they were successful at increasing freshwater survival.

Restoration actions should thus be viewed as experiments, where the intent is to determine, through manipulation of the ecological system and follow-up monitoring, which approaches work best in terms of achieving a stated management objective. Note that this recommendation applies equally to changes in fishing regulations, to habitat-oriented actions, and to socio-economic programs designed to aid restoration. Meeting the goals of thorough evaluation and accountability requires that restoration actions be

planned ahead of time to conform with rigorous experimental designs, which will permit evaluation of their relative effectiveness at some later date. Without such rigorous design, it will be difficult, if not impossible, to attribute some observed response in fish populations to a particular restoration action.

We review past experience and literature on the concepts of experimental design to develop a framework that will permit application of these concepts to restoration actions for salmon in the AYK region of Alaska. This document was written with two audiences in mind, those who will be undertaking restoration projects and those who will choose which projects to fund. This dual audience is consistent with an important theme of the AYK Sustainable Salmon Initiative (AYK SSI), which is to create a research and restoration program "... that is inclusive through a strong communication plan and public involvement ..." and "... is strongly committed to seeking a more inclusive process through capacity building" (AYK SSI 2004). Thus, to address the first audience, this paper provides a framework of concepts and components of experimental design that can be used by proponents of restoration projects (e.g. local or regional community groups, non-profit native corporations, other non-governmental organizations, and government agencies). By following this framework, restoration practitioners will be able to assess whether their project was able to meet its restoration objectives. The second audience, the AYK Scientific and Technical Committee (STC), will be writing a "Research and Restoration Plan" in the future. If experimental design criteria such as those described in this report are incorporated into that plan, it will result in the following benefits:

- The AYK STC or other funding groups will be able to evaluate proposed restoration projects based on whether they meet the requirements for good experimental design. As has been done in other restoration programs in western North America, that criterion will be an important component of funding decisions to permit evaluation of a restoration project's effectiveness.
- An implicit goal of restoration programs will more likely be met: learning more about the processes affecting AYK salmon populations.
- Accountability will increase at all levels and allow a thorough evaluation of the SSI restoration plans as a whole.
- The Arctic-Yukon-Kuskokwim Sustainable Salmon Initiative will be able to respond dynamically as understanding of regional ecosystems increases. Thus, the AYK SSI will be able to benefit from a more efficient use of its funds by focusing over the long term on the most effective restoration methods.

We recognize that in the last two years, some of the salmon populations that were a serious concern in the AYK region have recovered remarkably quickly, even without major interventions. This outcome is, of course, excellent for the salmon and people that depend on them, but it also clearly illustrates the need to conduct restoration projects (including changes in fishing regulations and other similar management actions) using experimental designs, as we describe in this document. For instance, imagine that millions of dollars had already been spent by the AYK SSI on various restoration efforts prior to 2001. If the same natural events had occurred as we have observed since then and if no provisions had been in place to monitor in an experimentally designed manner to account for the confounding effects of these events, many people would have falsely attributed the quick turnaround in salmon abundance to the restoration actions, creating unwarranted confidence in the ability of these actions to prevent or ameliorate future declines. Hence, an experimental approach to design of restoration projects is critical.

Despite the rapid recovery of several AYK salmon populations, the material that we present here is still relevant because the STC's future "Research and Restoration Plan" will be developing a framework that can be used not only for those AYK salmon populations that have not yet recovered, but also for any AYK populations that may be in need of restoration in the future. In addition, our material is also relevant in the context of precautionary actions, such as changes in management practices, that may be taken to increase the resilience of AYK populations and reduce the chance of future collapses.

In this report, we use the term "restoration (or recovery) actions, activities, plans, or projects" in a broad sense to reflect a wide range of human actions. These restoration actions, which could also be considered management "levers", potentially include changes in fisheries regulations, new management practices, alteration of freshwater habitats, various forms of enhancement, and socio-economic initiatives. These actions are intended to meet the objectives of:

- Restoring, or achieving recovery of, salmon populations to some higher level of abundance and productivity, or
- Maintaining salmon populations' resilience by removing stressors that might make the populations more vulnerable to future unfavorable conditions.

Structure of this paper

Our main objective in this document is to briefly review the extensive literature on experimental design of manipulations of complex natural systems, using examples from restoration of salmon populations. This information is relevant to restoration actions in the AYK region. After a brief Introduction, in section II we discuss several principles of good experimental design and develop a checklist of them (summarized in the Appendix). This checklist will assist practitioners who will apply the principles in their restoration work, as well as program administrators who must choose which actions to support financially. We also discuss some of the real-world challenges facing attempts to implement these principles and indicate approaches for dealing with these challenges. In section III, we discuss specific concerns regarding good experimental design and effective monitoring within the context of different types of restoration projects that might be considered in the AYK region, ranging from manipulations of freshwater habitat and use of incubation boxes to changing fishing regulations. In section IV, we focus on suggestions for coordinators in the AYK SSI, in various private and community groups who will be involved in restoration across the AYK region, and managers in the Alaska Department of Fish and Game (ADF&G). There, we discuss broad strategic planning of co-ordinated restoration initiatives, including recommendations for a broad data base of actions to track their effectiveness and help with future planning. We then conclude with some general recommendations, which are reproduced below in Box 1.

Box 1: General recommendations

The following recommendations apply not only to manipulations of salmon habitat in fresh water but also to management regulations that affect harvesting, and socio-economic programs aimed at salmon conservation.

1. All salmon restoration actions should have at least two clearly stated main objectives:
   - Achieve results in terms of salmon abundance (specifying which populations, over what period, etc.). This objective could include either increasing salmon abundance or preventing loss of resilience or further reduction in abundance.
   - Learn over time which restoration actions work best and which do not.
2. All restoration projects affecting a particular population should be co-ordinated to reduce the chance of unintended interactions between projects. This also applies to other activities that may affect salmon directly or indirectly.
3. For the same reason as #2, co-ordinate actions *across* salmon populations and species.
4. Restoration actions should be carefully planned as part of a larger experimental design to reduce chances that naturally occurring confounding factors such as simultaneous changes in climate conditions will make it difficult to determine the effect of these actions. This can be achieved by using treatments and controls.
5. A carefully designed monitoring and evaluation program should be implemented to determine over the long term whether the restoration actions had the intended effects. This recommendation includes not only data collection but also data analysis and dissemination of results; the latter steps are often not funded sufficiently.
6. It is also critical to monitor how well the restoration actions were actually implemented.
7. Restoration programs should address habitat-forming <u>processes</u>. There is a critical distinction between restoration actions that merely "rehabilitate" site-specific conditions (such as spawning areas, pools, banks, large woody debris, boulders etc.) and those that aim to restore the natural <u>processes</u> that form, rehabilitate, and maintain salmon habitat over the long term. An ideal goal for watershed restoration would thus be to re-establish natural watershed <u>processes</u> (Ebersole et al. 1997). If met, this goal would also reduce the need for human intervention in the future.
8. Restoration programs should follow similar guidelines as individual restoration projects to remain accountable and promote an institutional culture where learning is encouraged. In particular, institutions should avoid falling into a "sunk cost" trap, where initially promising approaches continue to be pursued even if they prove ineffective because (a) initial investments were high and there is a political need to amortize these investments, or (b) individuals or organizations have developed a vested interest in the continuation of their projects.
9. Restoration projects that do not follow a complete experimental design are not necessarily unacceptable, but compromises should be avoided if they sacrifice too much power to discriminate among hypotheses and thereby lead to uninformative designs. While it is prudent to implement only the minimal design necessary for evaluation, skimping initially on good design might result in large costs later.

Although these recommendations appear obvious and logical, it is surprising how often one or more of them is omitted in the rush to take some well-meaning action. In particular, there is often a lack of support for monitoring and evaluating results (recommendation #5 above). Because of our lack of complete understanding of salmon populations and their complex ecosystems, it is critical that we take the approach that unexpected results may, and probably will, arise from at least some attempts at restoring salmon populations. To separate the successful restoration actions from unsuccessful ones, those actions need to be implemented as part of carefully designed experiments and thorough monitoring must be implemented to evaluate performance of each of those actions.

I. Introduction

> "*The image of a broken chain helps explain how so much money could be spent on salmon restoration over the past couple of decades with such little result. Restoration programs have been fixing individual links in the life history chain. Little attention has been given to the restoration of whole life histories. A life history-habitat chain is a living system. It must be 100% complete - all the links habitable - or the system dies.*"
>
> - Jim Lichatowich (2002)

The substantial reduction in abundance of several species of salmon (*Oncorhynchus* species) in the Arctic-Yukon-Kuskokwim (AYK) region of Alaska during the last decade had a large effect on people in Western Alaska. These salmon populations have supported not only commercial fisheries but also subsistence fisheries that are important sources of food and culture for local residents. The AYK Sustainable Salmon Initiative (AYK SSI) was established in 2001 to address this problem. Among several key questions to be addressed by the AYK SSI are:

> "What caused these changes in salmon abundance?"
>
> "What can people do to reverse these trends?"
>
> *(paraphrased from AYK SSI 2004 and www.aykssi.org)*

Unfortunately, the answer to the first question is unclear. Many changes in the natural system and in human activities have occurred simultaneously in recent years, and it is generally not possible to completely separate the relative magnitude of contributions of those changes to regional trends in abundance of salmon species, which are king or chinook (*O. tshawytscha*), pink (*O. gorbuscha*), chum (*O. keta*), coho (*O. kisutch*), and sockeye salmon (*O. nerka*). To name just a few changes, there have been (1) reduced marine survival rates that were possibly due to climatic change or other processes reflected by coccolithophore blooms, (2) more beaver dams changing flow regimes and potentially blocking access to spawning areas as beavers expanded into the Norton Sound area over the last 20 years, (3) earlier break-up of sea ice and less of it in the Bering Sea (which reduces primary production at the base of the food chain in the spring, among other effects), and (4) changes in by-catch of salmon in the trawl fishery. Unfortunately (but as is typical), many of these changes have overlapped in time and space, which creates confounding among possible interpretations of the causes of the observed changes in salmon abundance. The interim report by the National Research Council (NRC) "Committee on Review of Arctic-Yukon-Kuskokwim (Western Alaska) Research and Restoration Plan for Salmon" came to a similar conclusion: "The committee judges that insufficient information is currently available to initiate a large-scale restoration program, although some small-scale local programs appear to be worth investigating" (NRC 2004). The NRC Committee therefore supports extensive research into the causes of decline in AYK salmon abundance as an initial priority.

Despite the lack of a clear answer to the first question about causes of reduction in salmon abundance, there are many suggestions offered by scientists, managers, and residents in the region to answer the second question about what we can do to reverse the trends in abundance. One approach would be to alter factors that appear to either limit the amount of spawning habitat (or its quality) or that reduce the survival rates of salmon at various points in their life history, including fishing.

However, experience with attempts to improve the status of drastically reduced salmon populations elsewhere on the North Pacific Rim has produced mixed results. There have been some clear successes, some dismal and expensive failures, and other cases where the outcome was unclear due to confounding factors or lack of rigorous evaluation (Williams et al. 1997; Stouder et al. 1997). Two examples illustrate some failures of restoration programs. First, several decades ago in the Pacific Northwest of the United States, streams through which salmon migrated and in which juvenile salmon reared were routinely cleared of downed trees and branches to "improve" the habitat (Sedell et al. 1988). However, subsequent scientific research demonstrated that at least some large woody debris is important to create the pools and other complex habitat structures that certain juvenile salmon require to find food and avoid predators (Bisson et al. 1987; Maser and Sedell 1994). A second example comes from the Yakima River sub-basin of the Columbia River in Washington State, where fish screens were installed at various irrigation and power diversions. Monitoring before and after the installation of screens showed that the juvenile survival rate roughly doubled for the spring chinook salmon smolts. However, the number of adults returning per spawner did not show any increase, possibly because of large variability in survival rates later in the freshwater or marine environments, which tended to mask or overwhelm the improved survival rate at the early life stage that had resulted from fish screens (Marmorek et al. 2004).

Some reasons for failures of past attempts at salmon restoration

Although the restoration methods used in the woody-debris and fish-screen examples are not applicable to the AYK area, these cases illustrate some generic problems that are relevant to all regions occupied by Pacific salmon. Both examples show that the dynamics and complex interactions among components of salmon systems are often not well understood, and some of those interactions can lead to counterintuitive or unexpected outcomes from well-intentioned human actions. The fish screen example in particular illustrates that fixing obvious problems is often not enough, even if the restoration action in question clearly succeeds at reducing mortality.

Of course, a crucial problem in this second case is that humans can alter reproductive success and survival of salmon in fresh water, but we have little influence over survival of fish in the ocean, other than by changing harvesting practices or harvest rates. This problem brings to mind the numerous cases of populations of wildlife and birds, as well as fish, for which increasing survival rate at an early life stage (e.g. through improving habitat or reducing predator populations that consume juveniles) has often been offset (or compensated for) to some extent by density-dependent growth or survival processes at a later life stage. As well, increases in early-life-stage survival rate can simply be swamped by much larger sources of variation in growth and survival at a later life stage, resulting in no noticeable increase in total abundance of adult salmon, as illustrated above by the case from the Yakima River.

These examples emphasize that in order to be successful, restoration efforts must be directed toward the key limiting factors that restrict the abundance or productivity of the salmon stock. Limiting factors are the processes or features of the physical or biological environment that constrain, or limit, abundance of a population of animals. In salmon populations, there can be various limiting factors on abundance of adult recruits, depending on the specific situation. For instance, in a watershed containing a sockeye salmon population, one can easily envisage a situation in which the quantity of good-quality spawning habitat might be reduced through road construction but in which all other habitats that the salmon occupy in later life stages remain relatively pristine and productive. In this case, taking action to increase the amount of good-quality spawning habitat will likely tend to increase the resulting abundance of fry and subsequent adults. Such situations are often referred to as "fry-recruitment limited" because the number of juveniles (fry) produced is the main limit on abundance of adults produced.

However, imagine a second situation, in which the rearing lake's productivity for that sockeye population deteriorated at about the same time as new road construction degraded the freshwater spawning habitat. In this case, an increase in fry abundance from improved spawning habitat might not result in a concomitant increase in adult sockeye abundance unless conditions were improved in the rearing lake as well. In short, survival rate of fry through to smolts migrating out of the lake will be the dominant limiting factor on abundance of adults, and will constrain their increase, despite substantial improvements in spawning habitat and production of fry. Hence, considerable attention must be paid to identifying the limiting factors that prevent recovery of salmon populations before spending funds on restoration actions. In this example with the degraded lake, restoration actions would be needed at both "bottlenecks" (spawning habitat and rearing lake) to have a reasonable chance of producing a substantial gain in adult sockeye salmon abundance.

Another way to view limiting factors is to consider Jim Lichatowich's view quoted at the start of this section. He considers a salmon population's entire life history from eggs to returning adults as an unbroken chain, where each physical link in the chain is analogous to a salmon life stage. An abundant salmon population will have all of its links in the chain intact, but a seriously depleted one may have one or more links "broken" (i.e. in the sense of functioning well below full potential, rather than not functioning at all). Successful restoration obviously involves identifying and fixing the key links and then maintaining all of them in the chain (i.e. over the entire life history of the salmon). Because more than one link may be "broken", it is a necessary, but not sufficient, condition for restoring a formerly abundant salmon population that a given vital link be restored to a fully functioning state. Other links may also need to be restored to achieve full recovery.

These are just a few of the reasons why so many salmon restoration actions have failed to live up to expectations. If we wish to avoid the mistakes of the past, it is instructive to recognize both these and other reasons for failure, as compiled in Box 2 (adapted from Marmorek et al. 2004 and Peterman and McAllister 1993).

Box 2: Reasons for failure of past actions to restore fish populations

1. *Unclear objectives*: The ultimate objectives of the restoration actions were unclear, leading to unfocused, counterproductive, or incomplete restoration actions.
2. *Poor implementation*: Restoration actions weren't actually implemented at all, or were implemented poorly.
3. *Degraded ecosystems*: The ecosystem was extremely degraded to begin with and therefore had little capacity to show a response to attempts at restoration.
4. *Limiting factor(s) not addressed*: Due to incomplete understanding of the complex interactions and feedbacks among processes affecting salmon population dynamics, restoration actions did not address the key processes limiting the biological response variable (e.g. abundance of adult salmon recruits), but rather addressed those processes that were easiest to manipulate.
5. *Actions too localized*: Restoration actions were too localized and did not affect the more appropriate, larger spatial scale.
6. *Too little time*: Restoration programs did not have enough time to create clear benefits to salmon survival rates before the programs were reviewed and cancelled. Here the true ecological effect of actions would be unknown, but the actions would be considered unsuccessful in terms of learning what works and what does not.
7. *Confounding factors*: Some concurrent changes in one or more confounding factors affected the fish population (e.g. drought, flood, changes in ocean productivity or harvest rate). This situation prevented scientists from uniquely attributing an observed change (or lack thereof) in a fish population to a particular restoration action. Incorrect conclusions were thus drawn about the actual effectiveness of the intervention, leading to a "failure" in the sense that either successful or unsuccessful actions may not have been correctly classified as such. This may happen in situations where adequate controls were lacking or statistical power was low (i.e., a small chance of correctly detecting some specified magnitude of an effect). Low power can arise from (a) large natural variability in fish responses to restoration methods, (b) observation error and bias in estimates of responses, (c) confounding by other concurrently changing factors, or (d) small sample size.
8. *Poor documentation*: Fisheries management agencies or other groups involved in restoration have not often documented either the reasons for past successes or failures, or relevant background data for judging why decisions were made (Larkin 1972). Therefore, it has often been difficult to pass the knowledge gained through experience from one group of scientists, managers, and community residents to the next (Hilborn 1992). This has perpetuated uncertainty about the potential effects of restoration actions. The result has been a persistent use of questionable methods, resulting in increased biological conservation concerns, foregone harvest, or potential overharvesting.

Experimental design and adaptive management

The principles that we advocate in this paper for designing and implementing restoration actions in the AYK region form the basis of adaptive management, in the formal sense as defined by Walters (1986) and Holling et al. (1978). Briefly, some common elements are to assess the current knowledge about the system, design some actions to be taken, implement them, monitor specific indicators of the system, evaluate the system's response to the actions, and adjust future actions according to what has been learned about the system's dynamic processes, as well as the system's response to previous actions. A key aspect of adaptive management is that there is an explicit value to the information that is gained from taking actions. We emphasize that this includes information gained from taking new management actions on fisheries, which is where the original concept of adaptive management emerged in applied ecology (Walters and Hilborn 1976). To permit effective learning, this goal of gaining information about how the manipulated system works must be one of the criteria for choosing actions, in addition to other, more traditional criteria such as maintaining or increasing harvests.

However, one of the well known problems with the application of adaptive management in ecological systems is that it often takes a long time to rigorously evaluate the effectiveness of some management action because of natural variability and unavoidable human errors in estimation of abundances of salmon. Depending on the situation and the life span of the key ecosystem components, a decade is not an unusual expectation for clear results. A notable exception is the large-scale experiment conducted on the groundfish community off the Northwest coast of Australia in the late 1980s (Sainsbury et al. 1997). There, various regions had different combinations of contrasting "treatments": trawling allowed or not, trap fishing allowed or not, and no fishing allowed. After five years of this experimental management regime, it was possible to clearly distinguish among four major alternative hypotheses about the mechanism that caused the decline of the commercially valuable species -- it was the destruction by bottom trawling of the large benthos that provided important habitat for the fish. In most ecological situations, though, results are not so clear-cut in such a short time.

It is important that people and agencies realize that such lags may be lengthy. Failure to do so will create unrealistic expectations and perhaps premature cancellation of promising approaches. An example of this might be the CALFED (California Federal Bay-Delta) Program, which has spent roughly $3 billion in the first four years of a 30-year program contributing toward restoring ecosystems in the San Francisco Bay-Sacramento/San Joaquin River Delta and upstream watersheds. However, CALFED has recently been criticized by the California government for failing to produce results that would justify the large expenditures. As a consequence, the State is planning to reduce CALFED’s budgets for the next fiscal year. Adaptive management is included in CALFED’s strategic planning documents, and agencies have made substantive efforts to co-ordinate research and restoration actions. However, more communication about the long-term nature of the project is clearly required.

II. Overview of Principles of Good Experimental Design

"*Each time history repeats itself, the price goes up.*" - Ronald Wright (2004)

Our main theme is that application of the principles of good experimental design to restoration actions in the AYK Sustainable Salmon Initiative will help to reduce (but of course, never completely eliminate) the chance that the problems in Box 2 will occur. Wise application of those principles will also contribute to a legacy of institutional and community-based knowledge about which restoration methods worked best and which did not, and in what situations. The failures are as important to know about, if not more important, than the successes because learning from failures means that less effort will be wasted on ineffective projects. We cannot overstate the important role of "learning by doing" as one of the broad goals of the AYK SSI. That goal should help shape and guide many activities of the AYK SSI, as should the related goal of aiming for "continuous improvement".

These goals are appropriate to consider now, even given the recommendation by the NRC committee that it is too early in our gathering of knowledge on AYK salmon systems to initiate a large-scale restoration program (NRC 2004). First, the criteria for such a restoration program will be written by the AYK STC in the near future. Second, to design future restoration projects well, key inputs will be required that can be provided by data collected by research that focuses now on the most critical questions concerning possible mechanisms causing changes in salmon abundance. In other words, if research and monitoring programs are extended soon beyond their current level, they could, with proper context, help to identify those mechanisms causing future declines, if such changes occur again. Finally, principles of experimental design are also useful for taking proactive measures that are intended to maintain resilience of salmon populations, even if intensive restoration should be deemed unnecessary at this point and a restoration program is deferred.

Considerable research and experience in the last few decades has led to an enormous amount of literature on principles of experimental design (e.g., Green 1979; Hurlbert 1984; Walters 1986; Hairston 1989; Eberhardt and Thomas 1991; Keeley and Walters 1994; Mellina and Hinch 1995; Schmitt and Osenberg 1996; Schwarz 1998). These principles are applicable to a wide range of activities, including environmental monitoring, environmental assessment, and experimental manipulations to the ecological system to achieve some management objective. Most of these principles can be applied to the specific case of designing restoration projects to strengthen the conclusions drawn from them (Walters and Green 1997; MacGregor et al. 2002). We draw upon this literature to extract some key ideas that are directly relevant to the design of restoration actions in the AYK SSI program.

Essentially, a key issue is how to choose the most appropriate experimental design for a restoration program given the numerous options for each principle of design. For instance, should the program have only one area treated with some restoration action, three, six, or ...? How many control areas should there be? For how many years should the experiment be conducted? How large should the treatment area and magnitude of human action be? These and other related questions can be answered in part by following some principles of experimental design.

A. Description of the principles of good experimental design

The following description of eight key principles of experimental design are summarized in the Appendix in a detailed checklist that is composed of questions that can be used by both planners of restoration projects and people who choose which projects to fund. These principles are drawn from the literature cited above. A major underlying purpose of applying these principles is to reduce the magnitude of the residual variation in the observed response variable that cannot be explained by the "treatment" that was used. In other words, if some restoration action is taken to increase the marine survival rate of a chinook salmon stock, for example, the experiment should be designed to minimize the number of potential factors that could confound the interpretation of the causes of subsequent changes (if any) in marine survival rate (e.g., factors affecting body size of chinook smolts, temporal trends in ocean conditions, competition with other stocks and species, etc.). Attempting to control for such confounding factors is done by using comparison groups (treatments, controls, contrasting treatments in space or time), replication, randomization, and stratification, as well as certain statistical methods of analysis (Figure 1). These ideas are elaborated in the following sections.

1. Clearly stated objective or hypothesis

The primary objective of a contemplated restoration action should be stated precisely and clearly. It is not advisable to leave an objective as a vague, broad statement such as "improve the abundance of chinook salmon in Norton Sound". Instead, it should be defined in specific, operational terms that identify measurable indicators that can be used to determine at some later date how well the objective has been met. For instance, a specific objective could read "Increase the abundance of adult recruits (i.e. catch plus escapement) of the Unalakleet chinook salmon population by 40% above last year's level within the next 18 years (about three generations)." Lack of such a specific objective makes it difficult to convincingly determine the success or failure of a restoration project. Such an omission would also create difficulties for effectively designing the treatment actions and a monitoring scheme, which may ultimately create one of the many problems listed above in Box 2.

Specifically, the particular objective determines which performance measures (indicators) can be used to determine success or failure. For instance, with the specific objective stated above for the Unalakleet chinook salmon population, it would be obvious that the evaluation of the success of the restoration project could not be based merely on comparing the abundance of chinook fry prior to the planned restoration action with that abundance afterwards (even if that was possible). Instead, total abundance of adult recruits for that system would need to be monitored, by year of spawning. Although this concept is simple, it is often overlooked. For instance, for many years, the Salmonid Enhancement Program (SEP) of the Canadian Department of Fisheries and Oceans used number of juveniles released from its various enhancement facilities as a measure of the success of SEP (Hilborn and Winton 1993), even though the goal was to increase the abundance of adult recruits, as well as harvests in commercial and recreational fisheries.

Another way to look at clarifying objectives is to think in terms of identifying hypotheses to be tested as part of restoration programs. In essence, the implicit hypothesis of most restoration programs is to test whether the restoration method will work to increase salmon abundance. For example, if a hypothesis is that spawner abundances for particular stocks are being limited by the by-catch in trawl fisheries, then various management restrictions on trawlers could be imposed to test that hypothesis, while at the same time attempting to increase abundance. Stock-specific identification of fish caught would be essential, though, to test this particular hypothesized link.

A different testable hypothesis of specific importance for restoration efforts concerns interactions among salmon stocks in the ocean. There is considerable evidence that the early marine life stage is the source of much of the year-to-year variation in recruits per spawner (Peterman 1987). There is also evidence that density-dependent interactions on growth and survival rates can occur between salmon populations and even between salmon species (Peterman 1984; Ruggerone et al. 2003), and early marine growth has in turn been positively correlated with marine survival rate in some stocks (Groot and Margolis 1991). A testable hypothesis could then be that "density-dependent interactions between pink and chum salmon in the ocean reduce marine survival rates of chums in the AYK region in years when numerous pink fry are present". Further elaboration on testing this hypothesis appears under Principle 3.

2. Scope of project (site, context, spatial and temporal scale)

Determining the appropriate scope for a restoration project is also an important element to achieving success. As used here, "scope" refers to:

- Specific sites on which restoration actions will occur (e.g., where will beaver dams be removed?)
- Spatial resolution, or how close each of the replicate restoration actions will be to one another (e.g., on adjacent streams or every third stream?)
- Spatial extent over which those actions will be spread (the full river basin, or just a section of it?)
- Temporal resolution, or how frequently the treatments will be applied (e.g. fertilization of a sockeye rearing lake once or twice per summer?)
- Temporal extent, or the time horizon over which restoration will be applied and evaluated through monitoring (e.g. applied over 5 years and evaluated over another 5 plus the time for two generations of fish, or ...?)
- Broad context in which the restoration actions are placed (e.g. are some of the beaver dams being removed on tributaries for which upstream spawning areas are already degraded due to siltation or other causes? As well, will the effects of manipulation of freshwater habitat on a particular stream potentially be confounded (either exaggerated or reduced) by restoration actions elsewhere in the watershed that affect fish from that stream later in their lives?)

Restoration planners should carefully consider these issues of scope. Failure to do so may induce various problems noted in Box 2, such as degraded ecosystems, actions too localized, too little time, and confounding factors. Problems with scoping are often a sign of lack of time devoted to planning, limited expertise, or insufficient background research, which also increase the chance of poor implementation and failing to address the actual limiting factors.

3. Use of treatments and controls, or contrasting treatments

Experimental units are the various periods, locations, fish populations, or other discrete entities for which data are collected on indicators and respective responses are compared. "Treated units" are those in which some experimental manipulation is done, such as reducing in-river harvest or removing barriers to upstream access. "Controls" are intended to be experimental units in which everything is the same as the "treated units", except for the experimental manipulation. These ideal features for control units are rarely met, because each location will be somewhat different than other locations, even if by only a small amount. However, searching for controls that are as similar to treated units as possible reduces the chance (but probably never completely eliminates it) that an observed difference between treatment and control areas is due to some intrinsic (but unrecognized) difference between the treatment and control areas, rather than to the experimental manipulation that was carried out in the treatment areas.

"Contrasting treatments" are another way to create comparison groups to determine the effect of some manipulation on the response variable being monitored. For example, some streams could be supplemented with 100,000 coho fry from incubation boxes, others could receive 300,000, and still others 500,000. If there were, for example, density-dependent effects on coho smolt production due to limited freshwater territories, one would expect them to appear more strongly in the latter experimental units.

The most important reason for using control units and contrasting treatments is to help confidently estimate the effect of the treatment (if any) compared to other possible explanations of the observed data. This use of experimental units that are treated differently creates an effective context for correctly interpreting system responses to restoration efforts, and thus helps to control for confounding factors such as temporal trends in ocean productivity that affect all experimental units.

To illustrate the benefit of comparison groups, consider the case of Norton Sound chum salmon. In the late 1990s, incubation boxes were used to increase egg-to-fry survival rate and fry were released along with wild fry. Adult abundance of chum salmon subsequently dropped, leading some people to conclude that the incubation boxes were a failure and others to conclude that the drop in abundance was merely a coincidence (Doug Molyneaux, ADF&G, Bethel, AK, personal communication). A clearer interpretation might have been possible if the treatment (i.e., use of incubation boxes) had been applied to a variety of chum salmon stocks simultaneously, while keeping an equivalent set of stocks free of treatment but monitoring them. If both treated and untreated stocks showed similar decreases in adult recruits produced per spawner or other response variable, then it would be clear that the observed decrease could not all be attributed to incubation boxes, suggesting, for instance, that reduced marine survival rate might have decreased for many regional stocks. Conversely, an observed increase in the response variable for treated stocks but not for the others would have led to the conclusion that using the treatment was having a positive effect and therefore would probably be worth doing again. A third situation might have been that only the treated stocks declined, due to some unknown mechanism. Thus, without such controls of untreated areas for comparison with treated areas, invalid conclusions could easily be drawn about the effectiveness of some restoration action from only observing treated units.

An example of an experimental design that effectively used a control unit is the fertilization of Salmon Lake in the Norton Sound district starting in the late 1990s and continuing through 2004, except for two years without fertilization. The abundance of sockeye salmon increased in both the Salmon Lake and nearby Glacial Lake (not-fertilized) populations, which indicates that some or all of the response in the Salmon Lake stock was likely due to factors other than fertilization, such as improved marine survival (Doug Molyneaux, personal communication).

To reduce the chances of natural variability in environmental factors confounding the results of an experiment, ecologists have applied several types of experimental designs. The simplest design is a before-after (BA) comparison at a single location. Data on a key indicator or response variable are collected at one time before a treatment begins and are compared with such data collected after the treatment, which essentially uses the pre-treatment period as a control. The biggest problem with this simple before-after design is that ecological systems are so variable that it frequently happens that some uncontrollable event (perhaps even unknown at the time) coincides with the start of the treatment or occurs during the treatment. This leads to ambiguity about the effect of the treatment, because the observed response, if any, may be attributable to the treatment, to the other uncontrollable event, a combination of the two, or even just a coincidence. Furthermore, before and after surveys at a single time each do not account for natural variability in responses.

To get around these problems, the BACI (before-after-control-impact) design (Green 1979) creates four groups for comparison: multiple (ideally) control sites monitored once both before and after treatment begins, and the same for treated sites. Then, if some confounding or coincidental factor changes during the treatment period, the effect of this change will show up in the untreated control sites as well as the treated ones, leading to the conclusion that the treatment was not the only source of the observed change in response variable. Although this is a big improvement over the simple before-after design, there is a good possibility that treatment effects will be obscured by large year-to-year variability in the response variable. To alleviate this problem, Stewart-Oaten et al. (1986) further extended the BACI design (BACI-P) to use data series collected simultaneously at several times in single or multiple control and impact sites, which creates paired time series for each. This design accounts for natural variability in the response variable before and after treatment in both control and impact sites. Underwood (1991) developed the enhanced BACI-P design to detect changes in the magnitude of variation in control versus treated sites. This design permits comparison not only of the mean response to treatment in control and impact sites, but also the variability in them.

Some biologists have expressed the view that density-dependent interactions between pink and chum salmon in the ocean may be affecting marine survival rates of chums in Norton Sound (Gene Sandone, ADF&G, Anchorage, personal communication). This is a testable hypothesis using a "natural" BACI-P experimental design, i.e. one without human intervention. In Norton Sound, even-year pink salmon abundances are much greater than odd-year pink runs, creating a large contrast between abundance of pink fry encountered by chum fry. In essence, every other year's data point for Norton Sound chum marine survival rates can be a replicate of a similar "treatment" (years with numerous pink fry competing for food but also providing a buffer from predation through the presence of many more prey, and years with few pink fry doing the opposite). This natural treatment and control situation is ideal because it is likely that other potentially confounding factors that could affect chum marine survival rates will not change regularly every other year, so if there is a strong "signal" from large pink fry populations, this should appear in the marine survival data for chums.

The BACI concept is a fundamental feature of good experimental design and can form a strong foundation for manipulations of management regulations or other activities aimed at restoring salmon populations. Note, however, that BACI designs may still be misleading if the sites to which treatments were applied are dissimilar in some important respect that has bearing on the response variable. Another related problem with BACI designs is that it is generally not clear to what degree results are transferable to other, similar sites elsewhere (Schwarz 1998). Replication helps to resolve both these issues, as noted in the next section.

4. Replication

If one wants to make sure that the results of an experiment are not distorted by hidden environmental factors, it is advisable to replicate both controls and treatments. In the simplest BACI design, one area is treated and one area serves as control. There, the difference observed between treated and untreated areas could be due to the treatment, but it could also be due to some inherent, but unnoticed, difference between the treated and control area. The more replicates of both treated and control areas there are (or the more replicates there are of each level of contrasting treatments), the lower the chance will be that an observed difference in average response between treated and control areas is due to some difference (other than the treatment) between those treated units and the controls (assuming that treatment areas are independent and randomly allocated, or systematically allocated to minimize the potential effect of environmental trends, i.e., using blocking - Box et al. 1978). Thus, replication helps to control for confounding factors. It can further help reveal these confounding factors, because if only one or a few treated units respond quite differently to a treatment, this would indicate that some other factor is affecting the two groups of treated areas differently. Also, by estimating the variability among replicate treated sites compared to variability among control units, replication indicates how repeatable responses are across units.

Therefore, another benefit of replication is that results of an experiment with a restoration method can be more confidently extrapolated to a wider range of situations than would be the case if no replicates were used. This is termed increasing the "universe of inference", or the range of cases for which we can infer the effect of similar treatments in similar situations (Walters and Green 1997). With little or no replication, scientists would be in danger of unjustifiably extrapolating results to other cases. More replicates would tend to reduce the chance of that inadvertent mistake. Such replication could thereby indirectly reduce the amount of repetition necessary in future studies.

Of course, no two locations in a natural ecosystem are identical and therefore the concept of "replicates" can best be described on a continuum, ranging from no similarity of two sites, to considerable similarity. The more similar the two sites are, the clearer will be the evaluations of the restoration actions (but also the narrower the universe of inference). Recent work on spatial scales of similarity among salmon populations is relevant to choosing replicates to control for changes in ocean conditions. Peterman et al. (1998), Pyper et al. (2001, 2002) have shown for sockeye, pink, and chum salmon that indices of productivity (adult recruits per spawner after accounting for within-stock density-dependent effects) tend to be positively correlated among stocks that have nearby locations of ocean-entry for their migrating juveniles. The closer those locations are, the higher the correlation (reaching a maximum of 0.55 for pink, 0.39 for chum, and 0.2 for sockeye salmon). Those within, say, 500 to 800 km tend to share early ocean conditions, which the evidence shows contribute strongly to variations in marine survival rates (Peterman et al. 1998; Pyper et al. 2001, 2002). Such stocks would thus be useful replicates to help control for changes in ocean conditions that would potentially confound evaluations of the success of restoration actions.

5. Randomization

Another principle of experimental design is to randomly choose ahead of time which of the useable experimental units are to be treated with the restoration method and which are to be used as controls. Randomization helps to reduce the chance that inadvertent biases in the selection and layout of treatments, controls, and replicates may preclude correct interpretation of results. Thus, randomization both reduces the chance of misinterpreting results due to confounding factors and increases the size of the "universe of inference" by removing some of the potential sources of erroneous conclusions (Schwarz 1998).

6. Sufficient sample size

One of the goals of experimental design is to determine the effect of some factor on a response variable of interest (e.g. the effect of a change in annual starting date of a fishery on the mean run timing of returns in future years). Of course, it is desirable to estimate that effect with considerable confidence. That confidence can be measured in two ways, either by the precision of the parameter estimate (i.e. how similar independent, successive estimates are to one another) or by the statistical power of the experiment (the probability, or chance, of finding a statistically significant effect if one actually exists at some specified effect size). [We omit discussion of a third measure here (bias) because in most cases, bias is unknown because the true value in nature is unknowable]. Both precision and power increase with increasing sample size and decreasing sample variance. Sample variance is reduced by applying the principles noted above where appropriate, and can be further controlled by stratifying samples into ecologically similar sampling units, as well as by using more precise (and hopefully less biased) sampling methods. These features of experimental design essentially attempt to "control for" potentially confounding factors to identify a clearer "signal" of the effect of interest amid the "noise" arising from other varying factors, including natural variability and observation error.

Statistical power analysis (Peterman 1990a,b) can be used to determine appropriate sample sizes for treatment and control units, or for each level of treatment in a contrasting treatment design. "Appropriate" is determined by whether there is sufficient statistical power to detect a pre-determined magnitude of response that is deemed biologically or economically important. For any given situation, there may be several feasible designs. Statistical power analysis is one way to help choose among alternative designs, such as 3 streams with beaver dams removed and 7 with dams left intact (design #1), the reverse (design #2), 5 streams in each category (design #3), or only 2 in each (design #4). The probability of correctly detecting the effect of removing beaver dams on chum salmon will likely differ considerably among these designs. The standard steps to include in statistical power analysis to determine appropriate sample size are illustrated in Figure 2. A key idea in this figure is that a variety of alternative experimental designs should be evaluated in order to find which design meets the desired power (or precision) with the smallest possible sample size. This will be judged, in part, by non-statistical criteria that are not part of statistical power analysis, namely economic, social, and other considerations (see section IIC below.

Several ways exist to estimate statistical power from a specified experimental design, and alternatively, determine the sample size required by a given design to achieve a desired power. There are standard methods covered in most moderately advanced statistics textbooks (e.g. Dixon and Massey 1983). Thomas and Krebs (1997) review power analysis software packages that use well-worked-out analytical methods that are associated with standard parametric tests such as *t*-tests. There are also power analysis packages available on the web. For unusual statistical models for which there is no analytical solution, Monte Carlo simulations can be run.

Power analysis requires some estimate of natural variation in the response variable (e.g. adult chums produced per spawner). If no data or educated guesses are available, a small-scale pilot study would be justified. Also a simulation study can be helpful to ‘test-drive’ project implementation, optimize a sampling scheme, and uncover potential problems (this also helps with problems from Box 2 such as having too little time for the experiment and confounding factors, but potentially others as well).

<u>7. Complete measurement of responses</u>

Together with selecting sites for doing restoration, selecting response variables or indicators and working out a protocol for how these indicators will be measured in the field is another critical step for creating an operational plan that meets the objective(s) of a restoration project. As noted previously, many interacting processes can influence the overall survival rate from eggs to adult salmon. If the objective is to increase the abundance of adult chum salmon and the associated catch, for example, then it would be incomplete to merely monitor and estimate the effect on survival rate to the fry stage of some manipulation in habitat. Claims of success of the restoration method based on such an incomplete measure of the response could not only be misleading, but could be a waste of funds if in reality such increases in freshwater survival rate did not translate through later life stages into increased adult abundance. Thus, response variables must be fully compatible with the scope and objectives of the restoration project in order to correctly determine the level of success of that project.

Another illustration of this point comes from the earlier hypothetical example in which some watersheds are "treated" with removal of beaver dams and others not. This, and any other small-scale experimental situation, will only be informative to the extent that individual adult recruits can be correctly classified according to the watershed where they were spawned. Since percentage harvest rates cannot be assumed to either be the same or strictly random across watersheds, this means that we must be able to identify such groups of fish in the catch, not just on the spawning grounds. If this level of discrimination is not possible, and recruits per spawner cannot be estimated for individual watersheds, then the entire set of watersheds with beaver dams removed would need to be treated as one experimental unit, rather than as several. Thus, the choice of the response variable must match the scale of the question being asked.

An additional criterion for selecting variables to monitor is the extent to which other restoration experiments will also be monitoring the same variables. This should be maximized. Substantial overlap of these variables across separate studies will create greater replication and stronger inference than would be possible if most studies measured response variables that were unique to each case. Thus, attempts to standardize monitoring protocols across similar AYK SSI restoration projects will greatly increase both the amount learned and the efficiency of limited budgets (see section IV below for elaboration on monitoring protocols).

In these ways, careful selection of indicators can help avoid some causes of failures in Box 2, such as having too little time for the experiment and confounding factors (and possibly also identify situations where there are degraded ecosystems, failures to address limiting factors, and too localized actions).

8. Monitoring

Two main types of monitoring should be done in conjunction with projects aimed at restoring previous high abundances of Pacific salmon (MacDonald et al. 1991).

- *"Monitoring of effectiveness"* estimates the response to the restoration actions. The primary purpose of effectiveness monitoring is to measure whether objectives are met by restoration. As noted above, such monitoring needs to be conducted for both treatment and control experimental units (or for all levels of contrasting treatments).
- *"Monitoring of implementation"* determines the extent to which the action's experimental design was actually implemented in the field. Often logistical constraints or misinterpretations can make the final implemented design different from the plan, especially in the case of implementing harvest regulations in fisheries, where non-compliance is sometimes an issue. Such a difference can diminish the value of a carefully planned restoration experiment. The smaller that difference, the more benefits can be reaped from using good experimental design at the outset of the restoration program. Hence, careful attention to the execution of technical details in an experimental design is critical. Note also that problems uncovered during implementation monitoring could help amend similar problems during implementation of comparable restoration methods elsewhere.

B. Summary of benefits of applying these principles

In summary, there are two key benefits of applying principles of experimental design. First, we are more likely to correctly determine in the shortest time which methods of restoring salmon populations work best. Second, an experimental design framework will use restoration budgets effectively by reducing funding of ineffective actions over time and focusing more on the successful methods. The alternative, a lack of experimental design, may perpetuate ineffective and costly management programs.

C. Challenges to applying these principles

In spite of the obvious advantage of applying these principles of experimental design while conducting ecological restoration, their application is less common than would be ideal, largely because of the challenges to implementation posed by real-world situations (McAllister and Peterman 1992a; Walters 1997; Houlahan 1998; Leschine et al. 2003; AMFSTP 2004). There are two primary challenges: technical constraints and institutional obstacles.

Technical constraints can arise from several sources. First, it is rarely possible to achieve an ideal experimental design. Compromises must be made due to logistical constraints of time, budget, availability of too few independent replicate or control areas, and other factors. However, such compromises will increase the chance of having confounded interpretation of results or decreased statistical power to detect specified effects of restoration actions, if they are present. Nevertheless, as detailed by McAllister and Peterman (1992), there are many past situations in which at least some elements of experimental design (such as contrasting treatments) could have been implemented if the people involved had been aware of the large benefits to implementing their actions as experiments (even non-ideal ones), rather than as actions that were "certain to work".

A second source of technical constraints on experimental design is the inherent natural variability of salmon ecosystems and our limited understanding of how such systems operate. Thus, Bunnell and Dunsworth (2004) describe the challenge as designing experiments to answer critical questions for which we have a sufficient understanding without ignoring other important hypotheses for which our understanding is incomplete. For instance, the extent and nature of interactions between different populations and species in the ocean is not well understood, which makes it difficult to allocate truly independent controls and treatments for restoration projects. The key to dealing with these types of uncertainties is to use best judgment to fill in knowledge gaps where necessary, but to consider the potential limitations arising from such uncertainties in future planning, analysis, and interpretation of monitoring results. Hobbs (2003) brings up a different, but frequently encountered challenge to learning -- the mismatch between the limited spatio-temporal extent of an experiment that is feasible to implement, and the much larger extent over which the outcome is then extrapolated to inform management decisions.

In addition to technical constraints to applying experimental designs, there are also institutional obstacles imposed by individuals and organizations that manage natural resources (Walters 1997; Houlahan 1998; McAllister and Peterman 1992; Leschine et al. 2004). A review of freshwater habitat restoration projects by the Adaptive Management

Forum Scientific and Technical Panel (AMFSTP 2004) in California summarized how four categories of institutional obstacles inhibited learning about the successes and failures of restoration projects. Even though these insights are drawn from CALFED's Ecosystem Restoration Program and the U.S. Fish and Wildlife Service's Anadromous Fish Restoration Program, they are broadly relevant to the design and implementation of restoration actions elsewhere, including the AYK region.

- Regulatory restrictions (e.g., endangered species legislation) may constrain the active manipulations or treatments that may be required to achieve the range of variation needed in an experimental design.
- Requirements stipulated by funding agencies may constrain the application of experimental design or limit its usefulness because:
    - objectives of different funding agencies conflict and may not identify experimental design and formal monitoring and evaluation as priorities,
    - the length of time required to complete and evaluate restoration experiments may not coincide with the timeline of funding cycles, or
    - projects with substantial monitoring expenses may be perceived as too costly and fail to receive sufficient funding.
- Staff involved in experimental approaches to management may be insufficiently trained, or existing staff may not have sufficient resources to properly store, analyze, synthesize, or disseminate relevant data – critical elements to experimental design and learning.
- Communication among project designers and implementers may be limited, thereby leading to poorly co-ordinated study designs and monitoring protocols, and restricting the potential level of inference.

Because it is unlikely that these technical and institutional obstacles will ever be completely resolved, it is important to recognize and address them when choosing among study designs. This will often require complex tradeoffs, since no single design is likely to be ideal in every respect. When considering tradeoffs among different experimental designs, note that designs that do not incorporate the principles laid out in section II to the fullest extent are not necessarily unacceptable, they just do not provide as much information, or do not allow learning as quickly or reliably as would be possible with a more rigorous design. Problems can arise however, when choices among study designs are not properly justified or documented, or when tradeoffs among options are not properly evaluated before a final design is chosen. The implications of leaving out certain design features need to be clearly considered and understood. For instance, there is little point in pursuing a project where the length of the evaluation period is dictated by political considerations instead of being derived from the properties of the ecological system and experimental design. When a design is diluted so far by conflicting demands as to be uninformative, or when external demands conflict with what a design can be reasonably expected to deliver, as in the case of CALFED (see box on Adaptive Management above), then project designers should not be surprised when the results do not provide insights into the effectiveness of restoration actions and the whole program is then perceived as ineffective.

Complex decisions and the associated trade-off evaluations are made easier when aided by formal approaches to compare alternative study designs. For instance, decision analysis (Keeney 1982; Peterman and Anderson 1999, Clemen 1996) provides a rigorous and quantitative approach to decision-making, while PrOACT (Hammond et al. 1999) provides a qualitative approach to exploring tradeoffs. Below we use an approach similar to PrOACT to illustrate how scientists and managers can improve choices among alternative designs of a restoration program. If desired, more quantitative decision-making frameworks can be used to guide selection among alternative experimental designs (e.g. Peterman and Antcliffe 1993; Keeley and Walters 1994; Walters and Green 1997; MacGregor et al. 2002; Parnell 2002).

The PrOACT framework uses a design table to summarize the relative advantages of different design choices and structures decisions (Table 1). This table has two dimensions. Across the top are the alternative study designs or options being considered and down the left side are the objectives, criteria, or attributes that a project designer cares about. A description of the objectives requires that the decision maker(s) be explicit about the criteria deemed important for the decision in question. At a minimum, a list of objectives for evaluating experimental designs should include some description of the learning potential of a study design (Objectives 1 through 3 in Table 1). That potential is determined by how well the design meets the criteria laid out in section II. We will be more likely to evaluate the project rigorously by meeting criteria such as a high level of experimental control (to reduce confounding effects), high degree of confidence in the results (high statistical power), and a moderate to large universe of inference. In addition, other equally important technical, institutional, management, and/or socio-economic evaluation criteria for choosing among restoration designs may include:

- *Ease of implementation, data analysis, or interpretation* due to limited organizational capacity or experience;
- *Financial cost* to implement the restoration design due to budget constraints;
- *Complexity* of co-ordination, collaboration, and sharing of resources due to the large number of agencies responsible for implementing a particular restoration action;
- *Required level of change from status quo* -- it may be difficult or undesirable to modify existing regulatory frameworks or alter existing monitoring and research programs; or
- *Socio-economic considerations* due to other management priorities (e.g., number of jobs created or constraints placed on other industries or activities).

Table 1. Design table for a beaver dam removal project. A * means that design is the best choice for a given criterion. The alternative designs are as follows:

- **Design A:** 10-yr BACI design with 15 replicates
- **Design B:** 5-yr BACI-P design with 6 paired sites each in 2 strata and 3 watersheds
- **Design C:** Retrospective before-after analysis for 1 site for which there are 20 years of pre-beaver spawner-recruit (SR) data and 10 yrs of post-beaver SR data
- **Design D:** Open beaver trapping in Norton Sound area and use opportunistic SR data to evaluate effect.

	Alternative designs of restoration experiment			
Components of objectives	Design A	Design B	Design C	Design D
(1) Level of experimental control	High	High*	Low	Low
(2) Degree of confidence in results	High	High*	Moderate	Low
(3) Universe of inference	Moderate	Large*	Small	Large*
(4) Ease of implementation, data analysis, and interpretation	Moderately difficult	Difficult	Easy*	Difficult
(5) Financial cost	Moderate	High	Very low*	Low
(6) Coordination, collaboration, and sharing of resources required	No*	Yes	No*	Yes
(7) Requires change from status quo	No*	No*	No*	Yes
(8) Other socio-economic considerations	No	Community-based project*	No	Benefits from trapping*

It is important that alternative designs are described explicitly enough that they can be evaluated in light of all the stipulated decision criteria. This can be done using various schemes, from simple yes/no checkmarks to relative and absolute rankings of how well a design meets each decision criterion. Managers and scientists can then use the design table to make informed and well-documented tradeoffs when choosing among design options. The challenge in making good choices is to decide upon those study designs that represent good compromises among a number of critical objectives. We recommend the following approaches to help make such informed tradeoffs:

- Drop alternative designs that do not meet minimum design requirements (e.g., low degree of confidence in results for design D in Table 1 might not meet requirements; small universe of inference might exclude option C, whereas high cost might exclude option B, which leaves option A as the only option that meets all criteria).
- Constrain choices to those that meet principles of good experimental design within some tolerable range (e.g., statistical power between 70% and 80% for detecting some effect size deemed important, such as a 30 % increase in abundance).

- Remove from consideration those criteria that are uninformative or that do not help distinguish the alternatives (e.g., information about statistical power and sample size may evaluate a design in a similar way and as both overlap somewhat).
- Assign rankings (either relative or absolute) and weightings to each objective in the design table to help select the 'best overall' design (e.g., option B in our example is ranked best for the most criteria).
- Establish a set of decision "rules" to facilitate tradeoffs (e.g., choose cheapest option subject to having minimum number of controls, minimum statistical power, etc. – if we assume that the winning option must score at least moderate on criteria 1-3, options C and D in Table 1 would be excluded, and option A would be selected over B because of its lower cost).

Clearly, application of a design table approach may not necessarily identify a full, ideal experimental design as the most preferable option, nor will it ensure that learning and evaluation objectives can be met in every case. It will, however, help avoid bad compromises, which may result when specific objectives (e.g., 1-3, Table 1) are below some minimum requirement. Designs that do not meet such minimum requirements tend to spread effort out too thinly and haphazardly, provide uncertain or inaccurate conclusions, or have limited relevance to other circumstances. If for some reason it is impossible or undesirable to use a more informative design, the most appropriate choice may be no formal data collection and analysis at all (with the exception of minimal implementation monitoring). Alternatively, if ecological, financial, or other types of uncertainties and risks are too high to proceed without monitoring and evaluation, it may be best to choose a different restoration method.

III. Types of Restoration Actions: Problems and Solutions

In this section, we discuss the above principles of good experimental design and effective monitoring within the context of different types of restoration actions, or actions closely related to restoration efforts (Box 3). For each type of action, we discuss some of the characteristic problems encountered and potential solutions.

Box 3: Restoration actions

The list below includes general types of restoration actions, or management "levers", that have been used in various regions inhabited by Pacific salmon. Although some categories are not directly relevant to the AYK region, given its relatively pristine freshwater habitat, all categories are pertinent to the extent that they provide useful general lessons about attempts at restoration elsewhere. With the exception of the last category, we describe examples of these actions in section III of this report.

1. Collect baseline data to determine whether restoration is needed, what the most appropriate options are for intervention, and to compare with data collected after restoration takes place.
2. Change fisheries regulations and management practices (e.g. reduce overall harvest rates, reduce by-catch of salmon in various fisheries, change the allowed timing and location of fishing, alter mesh sizes on nets, etc.).
3. Restore freshwater habitat by altering factors that affect the amount and/or quality of that habitat, e.g.:
   - Remove barriers to upstream or downstream migration (e.g. beaver dams);
   - Improve or rehabilitate physical habitat after logging and mining;
   - Restore salmon habitat-forming processes. For example:
     - -- Remove artificial structures to permit lateral migration of stream channels.
     - -- Restore manipulated flows to mimic the timing, duration, and magnitude of the natural flow regime.
     - -- Maintain or restore natural vegetation cover throughout the watershed.
4. Artificially enhance salmon productivity.
   - Supplement fry abundance using hatcheries or incubation boxes, create new spawning habitat, remove predators on salmon.
   - Fertilize rearing lakes for sockeye salmon.
   - Fertilize rivers to make up for nutrient deficits due to low spawner abundances.
5. Initiate socio-economic programs.
   - Educate and facilitate communication of traditional knowledge.
   - Foster community stewardship and empower communities to take an active role in restoration.
   - Provide incentives to encourage conservation; remove incentives that encourage overexploitation and habitat destruction.
   - Consider certification or other forms of value-added marketing.

6. Ensure that technology, equipment, and personnel skills are adequate for addressing restoration needs by:
   - Budgeting for testing and training, and
   - Monitoring the implementation of restoration projects.
7. Use best management practices (BMPs) to reduce cumulative negative watershed effects of many small destructive actions such as road construction and/or removal of vegetation in riparian areas, which might result in sedimentation of streams, altered drainage patterns, and loss of riparian buffer zones.

A. Collection of baseline data

Although the collection of good baseline data on the environment, salmon, harvesting activities, and human activities is not considered restoration *per se*, it is a critical requirement for the design and evaluation of restoration actions in several respects (Rose and Smith 1992; Bisbal 2001; Bash and Ryan 2002). Baseline data provide:

- our primary means for tracking changes in salmon abundance and alerting us to declines that might necessitate restoration actions;
- long-term, systematically collected data series that are critical for understanding and addressing the full spectrum of ecosystem interactions, including long-term and time-lag effects;
- data for untreated time periods/areas/populations that can be compared with data from restoration projects to evaluate their effectiveness;
- often the only opportunity for retrospective evaluation of restoration projects and other management actions for which no specific pre- or post-treatment data were collected,
- a base against which to compare the outcomes from natural or unplanned experiments, such as a closure of an interception fishery, a sudden increase in the abundance of a key predator, or a year with exceptional ocean conditions;
- a historical context for setting restoration targets (e.g., historical distribution and abundance of salmon populations -- Stouder et al. 1997 and the range of natural variability in environmental conditions -- Morgan et al. 1994; Landres et al. 1999; Fowler and Hobbs 2002).

In effect, collection of baseline data can be seen as making deposits into a contingency fund for future scientific and management experiments. These deposits can then be accessed immediately as the need arises, without having to invest time and money in pre-treatment monitoring to establish a baseline for evaluation. Since time can be critical in saving declining stocks and managers can rarely afford to delay taking action until pre-treatment data are collected, deposits into the scientific contingency fund in the form of baseline data are a wise investment, even if such efforts confer little prestige or immediate returns on investments at the time. While it is possible to "substitute space for time" and draw conclusions about the effectiveness of a restoration action by solely relying on data from nearby untreated areas as a reference, this substitution usually results in relatively weak conclusions, for the same reason that simple before-after comparisons are fraught with pitfalls (Schwarz 1998), as discussed in section IIA3.

While a good baseline monitoring program is the only way to provide much-needed historical context, there are some caveats that should be observed when using baseline data. In particular, one must assure that the scale and nature of the data are appropriate for the needs of the particular restoration project at hand. For example, estimates of returning recruits are derived from catch and escapement data, which provide a fairly good estimate of gross returns for a stock or basin, but are not necessarily representative of returns to individual watersheds. Therefore, these data are of limited use for evaluating the effect of small-scale experiments that use individual watersheds as treatment units. Another caveat concerns the use of historical data to set management targets or predict ecosystem response to human manipulations or global change. As noted by Hilborn and Walters (1981), past observations may not accurately characterize responses to "extreme events", i.e., low probability situations that can affect the system in nonlinear or counterintuitive ways. This issue is clearly relevant to the AYK region because climatic change is already creating conditions that have rarely, if ever, been seen before. As well, human disturbances are often more severe than natural disturbance regimes (i.e., the frequency, magnitude, and scale of human disturbances can be outside normal ranges).

Baseline data collection programs can be made more effective as a 'scientific contingency fund' if standard monitoring programs are augmented with additional variables that will allow rigorous testing of hypotheses about sources of variation in salmon abundance and productivity in the future. This would essentially extend a baseline data collection program by adding a scientific research component, thereby promoting our understanding of the system and enabling future managers and restoration practitioners who need to take action to aid recovery of depressed stocks to quickly launch the types of restoration actions likely to be most effective in reversing declines in spawner abundance. The AYK Sustainable Salmon Initiative has a great opportunity to do just this because the recent recovery of many of the region's salmon populations has likely taken away some of the pressure to implement immediate recovery actions, thereby freeing funds for proactive projects that increase our general understanding of the system and help to guard against, or prepare for, future problems. This would be a good opportunity to fill any gaps in existing monitoring programs and establish a good quality baseline of stock assessment and environmental data.

In this context, for such a program to be effective in terms of promoting our understanding of the causes of change in adult salmon abundance and productivity, we recommend that high priority be placed on establishing as many programs as possible for estimating annual abundances of fry or smolts, in addition to stock-specific age-structured data on abundance of adult catch and spawners. Such data will permit separate annual estimates of stock-specific survival rates in the freshwater and marine life stages. This will provide valuable information to narrow down where to look for possible causes of a decline in salmon abundance in the future -- in the ocean or fresh water.

As with design of ecological and management experiments, there are some basic design questions that need to be answered when designing a baseline monitoring program, such as where, when, and how often to collect data, how to maintain compatibility with existing data series, and how to balance the various demands placed on a baseline monitoring program. A discussion of these issues is beyond the scope of this paper. However, there are some good examples of well co-ordinated baseline monitoring programs and standardized data collection in ecological systems that address these issues in some detail:

- U.S. Environmental Protection Agency's EMAP (Environmental Monitoring and Assessment Program) provides a standardized protocol for collecting many types of baseline data (http://www.epa.gov/emap/);
- Washington Salmon Recovery Funding Board (http://www.iac.wa.gov/srfb/docs.htm) and California Department of Fish and Game (Collins 2003) use standardized habitat monitoring protocols to monitor effectiveness of freshwater restoration;
- Collaborative Systemwide Monitoring and Evaluation Project (CSMEP) in the Columbia River Basin sets out monitoring protocols (Parnell et al. 2005).

B. Changes in fisheries regulations and management practices

The main actions that can be taken to influence abundance of salmon in the AYK region are those related to harvesting and escapement to spawning grounds. Closures to commercial, sport, or subsistence fishing for salmon in various times and places are the main regulatory method, along with restrictions to allowable fishing gear. Increased escapement resulting from various changes in regulations will obviously help to rebuild low-abundance salmon populations over time in "fry-recruitment-limited" systems, as long as there are no other density-dependent "bottlenecks". Of course, in terms of employment and food, there are substantial disadvantages of reducing fishing in the AYK region, so it is essential to rigorously evaluate the effectiveness of management regulations at increasing salmon abundance in the long term. In this context, another method to decrease harvest (but certainly a less desirable method in terms of life style and culture) is to simply "buy out" commercial harvesters for a few years, and, if it proves necessary to limit subsistence harvest as well, to provide alternatives to local salmon until stocks have recovered.

Unfortunately, in fisheries around the world, management regulations usually are not considered as part of an experiment. Agencies apply what they consider to be the best regulations given the circumstances and do a relatively straightforward before-and-after comparison. However, as noted above, such comparisons are a relatively weak experimental design that has often made it impossible to distinguish between alternative hypotheses about the causes of observed changes (various natural events or regulations?) (Walters 1986; Walters et al. 1988; McAllister and Peterman 1992a).

Fisheries management agencies obviously face considerable technical and institutional limitations to applying any but the simplest experimental designs. Independent replicate populations of both fish and harvester groups may be quite limited. Political considerations, special interests, and legal constraints may interfere with plans for a good design and with completion of a multi-year experiment. Effects of regulations are also quite susceptible to various confounding influences throughout the physical, ecological, and socio-economic systems.

Successfully applying an experimental design for management actions will reduce the influence of confounding factors and thus increase the chance of clearly identifying the effect of changes in regulations. In general, many fisheries management regulations could be implemented as part of experimental designs, akin to the adaptive management approach described above (Collie et al. 1990). This approach might be applicable to time and area closures in different places in alternating blocks of years, different restrictions on fishing gear, etc. For example, imagine that new regulations on mesh size are being proposed to reduce the by-catch of AYK salmon in trawl fisheries in the Bering Sea. The standard way of implementing the new regulations would be to have all vessels in the relevant trawl fleet in the region adopt the new mesh size. However, the by-catch rate might subsequently increase or decrease due to factors not related to mesh size (as occurred in a similar situation in the Fraser River sockeye fishery that caught chinook salmon as by-catch). That confounding would preclude determining how well the change in mesh size worked. If instead, the new mesh size regulations were used by half the trawl fleet, while the other half used the old mesh size (creating a BACI design), then it would be more feasible to determine the effect of the change in mesh size on the by-catch rate of salmon, even if independent but simultaneous changes occurred in the abundance of salmon or their catchability.

A key to implementing such a design is to establish a profit-sharing or other arrangement that will not put vessels using one type of fishing gear at a disadvantage compared to vessels using the other gear. This has been done before in British Columbia and would be most successful in the AYK if the trawl fishing industry is brought into discussions of how to design and execute the experimental evaluation of methods to reduce salmon by-catch. Objectives could also be met through altering harvest levels via various economic incentives, such as a bonus for consistently low by-catch rates.

Monitoring of the implementation of fishing regulations is a very important aspect of fisheries management because of the incentives that some harvesters and processors have to attempt to get around regulations. Transponders and other remote surveillance technologies, as well as on-board observers, can facilitate monitoring of certain types of regulations on commercial vessels, such as area and time closures and discarding limits. Monitoring and enforcement can benefit from including stakeholders in the design of experiments with management regulations. The development and testing of different enforcement options can also be done in the context of the principles of experimental design described here (see section IIIF below).

Because actions by fisheries management agencies both affect, and can be influenced by, the effect of other restoration actions, planned management actions should be integrated with other restoration projects where feasible. For instance, in some situations it may be possible to utilize a range of harvest levels as contrasting treatments to evaluate how freshwater restoration actions or enhancement projects respond to a wide range of spawner abundances. To alleviate some of the problems above and to identify the most promising designs, computer simulation models can carry out a virtual "pilot project" for such experiments before trying them out in the field.

C. Restoration of freshwater habitat

In temperate regions of western North America, stream manipulations are commonly used to alter environmental conditions in freshwater habitats so that they are more favorable for Pacific salmon, either by increasing the productive capacity of habitats or improving survival of early life stages. These actions can be grouped into four general categories (a) alterations to the physical structure of stream habitats (e.g., additions of large woody debris, placement of boulders, creation of pool habitat), (b) improvements to habitat connectivity and fish passage (e.g., installation of fish ladders, removal of natural or artificial barriers such as beaver dams or perched culverts, (c) recovery and maintenance of riparian areas (e.g., riparian planting, exclusion fencing, bank protection), or (d) restoration of upslope watershed areas away from streams (e.g., slope stabilization, road improvements) (Slaney and Martin 1997; Roni et al. 2002). Only a few of these actions, however, are directly relevant to the AYK region now or in the near future. Some of those might include actions to mitigate or offset effects of logging, mining, and related road construction activities, or improve habitat connectivity by removing barriers created by beaver dams or road crossings.

Poor design and monitoring of habitat manipulations constrain learning

In spite of substantial financial investment in freshwater restoration actions such as these, a common criticism is that small-scale projects at the level of stream reaches a few hundred meters long fail to provide sufficient insights into their potential effectiveness at broader scales (Kershner 1997; Bayley 2002; Roni et al. 2002; Shields et al. 2003; Bernhardt et al. 2005). This shortcoming has been attributed, in part, to poor experimental design and lack of sufficient monitoring of restoration projects.

Three spatial scales have been used in the design and implementation of restoration projects, each of which represents a different scope of project and affects learning about effectiveness of restoration in a different way. First, an individual project may concentrate actions in a single stream reach but lack coordination in design, implementation, and monitoring with other restoration actions, additional human activities, or connected habitats in a watershed. Such projects may be easy to implement, but insights may be limited to only whether implementation was effective at that particular location or whether local conditions changed, rather than providing insights at a broader scale. For instance, monitoring of an individual project to rehabilitate a stream after placer mining may detect increases in juvenile chinook salmon abundance associated with that action, but would not be able to indicate whether that change in abundance at that location was due to the improved habitat or simply due to shifts in patterns of habitat use (e.g., movement of juveniles away from some other area to the newly rehabilitated habitat). More generally, watershed connectivity will always pose problems when trying to manage one part in isolation from the others. For instance, any positive or negative developments elsewhere within the watershed may confound the effect of a project. Individual isolated projects may also not be able to identify whether critical limiting factors or bottlenecks in habitat production were affected by the project (e.g., actions to improve summer rearing habitats for coho salmon may be ineffective if over-wintering habitats are the main constraint on juvenile survival). An exception of course, would be if a restoration practitioner had some indication, by means other than a full watershed assessment, that only one key modification was required (e.g., removal of a beaver dam or perched culvert).

A second, often preferable approach to designing a restoration project would be to co-ordinate a number of projects in a single watershed, thereby leading to more effective restoration or a better understanding of the main constraints to local fish production. This type of co-ordination is commonly implemented for studies that emphasize restoration of key habitats and processes that aim to increase the amount and quality of freshwater habitats. Similar, but less typical examples include intensively monitored watersheds (i.e., long-term watershed studies) that are research-oriented and emphasize learning about the effect of human activities on fish and their habitats (e.g., Carnation Creek, B.C. -- Hartman et al. 1996, and Fish Creek, Oregon -- Reeves et al. 1997). Such approaches allow for a better assessment of limiting factors in a watershed and ranking of restoration needs (Roni et al. 2002; Beechie et al. 2003) than projects implemented in isolation. As well, such research-intensive watershed studies may help to identify reasons for local restoration successes or failures by providing data on how conditions in an individual watershed responded to a suite of actions. However, these observations may not necessarily be extrapolated to larger spatial scales due to a lack of replication and contrasting restoration treatments across different watersheds. Also, the results from such single-watershed designs cannot indicate whether some broader-scale factor was driving an observed change. For instance, changes in climatic variables, such as air temperature and precipitation, may have affected water temperatures and discharge, respectively, thereby leading to changes in juvenile survival rates and abundance. Lack of statistical independence among replicates may also be a concern for watershed studies because typical spatial scales, such as microhabitats, reaches, and streams, represent a nested hierarchy of spatial scales and varying magnitudes of connectivity among habitats (Frissell et al. 1986).

The third, and probably best design from a learning perspective co-ordinates numerous actions across many watersheds, either as part of a replicated experiment (e.g., Roni and Quinn 2001), or within a broader restoration program. Examples of the latter include British Columbia’s Watershed Restoration Program (Slaney and Martin 1997), CALFED's Ecosystem Restoration Program (http://calwater.ca.gov/Programs/EcosystemRestoration/Ecosystem.shtml), and the Anadromous Fish Restoration Program (http://www.delta.dfg.ca.gov/afrp/) AMFSTP 2004). Compared to the other two spatial scales, these types of studies have the greatest potential to detect statistically significant effects and to reduce confounding influences because they allow for spatial and temporal replication, use of controls, and contrasting treatments through space and time (Mellina and Hinch 1995), (e.g., staircase design of Walters et al. 1988). However, these designs may not be broadly applicable in the AYK context. Because few AYK watersheds require restoration and the physical and biological characteristics of watersheds and stream reaches are often unique, it is unlikely that there are enough reasonably similar sites or watersheds in need of treatment to use as replicates. A further challenge is that with an increased number of projects comes increased institutional and logistical constraints, i.e., complexity in co-ordinating actions, ensuring compatible data collection, and standardizing monitoring methods among projects across a broad geographic area (Bernhardt et al. 2005).

Unanticipated responses to restoration may result from limited understanding

As the following examples illustrate, our incomplete knowledge of the interaction among restoration actions, watershed processes, stream habitats, and salmonid populations has often hampered habitat restoration. For example, an incomplete understanding of interactions among various physical and biological components of watersheds has led to counterintuitive restoration outcomes. As mentioned earlier, the lack of knowledge about the beneficial role of large woody debris (LWD) resulted in its removal for decades. The intention was that this action would improve adult fish passage, but unintentionally resulted in the degradation of rearing conditions for juvenile salmon.

More recently, there has been a shift in perspectives about approaches to stream restoration (Hillman and Brierley 2005). The paradigm is changing from an engineering perspective, which focuses on rehabilitation of localized stream habitats (e.g., to stabilize eroding stream banks with riprap and riparian planting) to an ecological perspective, which attempts to restore habitat-forming processes in a watershed (e.g., allowing natural channel migration across a sufficiently vegetated valley floor). Our failure to recognize the importance of restoring such natural processes as part of past restoration projects may have created a legacy of engineered structures that can be costly to maintain and may ultimately be unnecessary or even counter-productive.

Incomplete understanding of ecological linkages may be of special concern in the AYK region, because its northern freshwater environments are different from temperate regions where most restoration research has occurred and the AYK habitats may not respond in the same way as in those other locations. For instance, Mossop and Bradford (2004) found in streams from the upper Yukon River that the abundance of pieces of large woody debris per unit of streambed was similar to that reported in the Pacific Northwest, but the size and total volume of woody debris were much lower. Due to the slower growth rate of trees in the Yukon, it may therefore take much longer for riparian trees to grow to the stage where they provide woody debris for nearby streams. Thus, human activities in riparian areas in the northern boreal forest in the AYK region may have a larger effect on the availability of woody debris in stream habitats longer than in temperate regions.

The removal of beaver dams is an example relevant to the AYK region for which there may be an incomplete understanding about the effects of restoration actions on salmonids, requiring an explicit description and testing of several clear hypotheses. Imagine a hypothetical situation in which restoration practitioners in Norton Sound inventory the number and size of beaver ponds on salmon-bearing streams and find that some action is warranted. Based on observations in the Pacific Northwest (Swanson 1991), they assume *a priori* that a reduction in beaver activity (by removal of dams and trapping of beavers) would improve fish passage for some species (e.g. downstream-migration chum salmon fry) but reduce over-wintering or rearing habitat for other species (e.g. coho salmon). Because monitoring of biological variables can be expensive, they design a study to only test which of three restoration techniques (with varied costs) is the most effective at reducing beaver activity over 5 years. However, at the end of the study, escapement surveys indicate that abundances of both chum and coho salmon are declining in treated streams. This leads to a subsequent study, which after another 3 years identifies that beaver ponds in the area act as sediment traps, thereby reducing sedimentation of salmon redds and improving egg-to-fry survival rates in downstream spawning areas – yet another mechanism that was initially overlooked.

In this hypothetical example, an incomplete understanding about ecosystem processes and failure to consider alternative, but plausible, hypotheses led to an unanticipated outcome. Thus, monitoring programs should be designed broadly enough to address ecological hypotheses, and these hypotheses must be cast widely enough to cover the full range of possible interactions. In addition, attempting to keep project costs to a minimum might result in greater expenses later. Thus, restoration actions should be applied cautiously in the AYK region, with high priority placed on learning, which can be based on good experimental designs and rigorous evaluations.

D. Artificial enhancement

The term enhancement has been used to cover a wide range of activities intended to increase salmon abundance and productivity (Gardner et al. 2004). Here, we are primarily concerned with practices that employ some form of intensive management akin to techniques traditionally used in agriculture, notably river and lake fertilization, spawning channels, incubation boxes, hatcheries, sea ranching, and genetic enhancement. In general, the design and evaluation of enhancement projects poses difficulties very similar to those associated with the design of freshwater habitat restoration projects. However, the potential for interaction between enhanced and wild populations poses some additional challenges.

Uncertain benefits and risks suggest monitoring and a focus on learning

In spite of the long-standing and widespread application of hatcheries and other enhancement technologies, scientific understanding of the effects and effectiveness of enhancement is still contentious. Benefits of enhancement to natural populations may be substantial (Brannon et al. 2004), as may be the risks (e.g., Levin et al. 2001; ISAB 2002; Weber and Fausch 2003; Gardner et al. 2004; Myers et al. 2004). For instance, in Canada and the U.S. Pacific Northwest, attempts to supplement wild populations of salmon with fish reared in hatcheries has resulted in frequent, but not uniformly occurring, problems such as spread of diseases and higher harvest rates that deplete less productive wild populations. Experience shows that both positive and negative impacts of enhancement technologies tend to be very site-specific and vary widely across projects, making it difficult to generalize observations from individual studies and establish generally applicable protocols and guidelines (e.g., Gardner et al. 2004).

The high degree of uncertainty and risk associated with salmon enhancement implies that projects should at least initially focus on learning, within a carefully designed framework that maximizes scientific inference through experimental design while limiting risks to wild salmon populations (Hilborn 1999). Enhancement-free zones comprising complete river systems and associated near-shore habitat used by juveniles and spawners may be helpful, because they can serve as experimental controls as well as preserving genetically and physiologically healthy populations (Hard 2002, Routledge 2002). However, the high degree of sensitivity of salmon response to site-specific factors means that comparisons between geographically distinct areas (i.e., using spatial replication and control) must be used with caution (Hilborn and Winton 1993) and should be augmented by before-after comparisons.

When structuring a monitoring program and selecting indicators, it is important to pay attention to not just local, short-term impacts, but also the broader context, including environmental factors, the scale of production relative to natural population, and longer-term lagged and/or cumulative effects (Gardner et al. 2004). Summaries such as Table 2 below, which identify the key risk factors associated with different restoration and enhancement activities, can be useful to make an initial assessment of the risks posed by a particular project, and can help in guiding experimental design and selection of monitoring endpoints.

Table 2: Relative risk to wild fish posed by various enhancement methods (adapted from table 8 in Gardner et al. 2004)**.** Values shown reflect estimated average risk for enhancement projects that are 'large' in terms of production. Smaller projects carry proportionally less risk, but effects of multiple projects should be expected to be cumulative.

Interactions	Obstruction Removal	Habitat Restoration	Lake/Stream Enrichement	Channels	Controlled Spawning	Controlled Rearing	Hatcheries
Mixed Fisheries	unlikely	none	possible	likely	likely	likely	likely
Genetic	unlikely	none	unlikely	possible	likely	likely	likely
Competition	none	possible	possible	likely	likely	likely	likely
Predation	unlikely	possible	possible	possible	possible	possible	possible
Fish Health	none	none	unlikely	possible	possible	possible	possible
Hab. Degradation	none	none	none	possible	possible	possible	possible

<u>Hatchery/wild interactions are subject to confounding environmental influences</u>

Like evaluation of other restoration actions, evaluation of enhancement projects is hampered by the considerable natural variability in vital rates and abundances of salmon populations. For example, a natural increase in sockeye abundances made it difficult to conclusively evaluate the effectiveness of the Salmon Lake fertilization project in Norton Sound District described in section IIA3. Similarly, difficulties in separating enhancement effects from natural variation in ocean survival rates have thwarted conclusive evaluation of both the British Columbia Salmon Enhancement Program (Hilborn and Winton 1993, Wood 2002) and the Prince William Sound (PWS), hatcheries in Alaska. The latter case has sparked considerable debate in the scientific literature (Eggers et al. 1991, Hilborn and Eggers 2000, 2001, Wertheimer et al. 2001, Wertheimer and Heard 2002), with some authors arguing that enhancement has increased overall production, but others claiming that it has merely supplanted natural production.

The influence of environmental factors on the outcome of enhancement projects is further complicated by the fact that interactions may be non-linear and may only manifest themselves under extreme conditions, such as in years with unusual food distribution or composition, or in years with extremely high or low abundances or survival rates. For example, Levin et al. (2001) found "a strong, negative relationship between survival of [wild Snake River spring] chinook salmon and the hatchery fish released, particularly during years of poor ocean conditions". Thus, even if we assume that the PWS hatchery program has indeed succeeded at increasing overall PWS salmon production, it is not clear whether such increases can be sustained during a period of unfavorable environmental conditions, when ocean carrying capacity may be substantially reduced (see also Peterman 1991; Hilborn and Eggers 2000).

To account for the effect of environmental influences in the assessment of enhancement projects, longer periods of observation than are common for management experiments may be required. Hilborn and Winton (1993) suggest 50 years as an appropriate time frame, though it may be possible to reduce this period by using an experimental setup specifically designed to address environmental variability. Such a setup may include a BACI-type experiment with enhancement-free zones as formal spatial controls. To increase power and distinguish between treatment and environmental effects sooner, we can allocate treatments in a staircase design (Walters et al. 1988), or combine studies specifically targeted towards gaining a better understanding of links between ocean conditions and survival of wild and enhanced fish with statistical methods that are suitable for non-replicated or poorly replicated whole-ecosystem experiments (Carpenter et al. 1989).

Other restoration actions may confound evaluation of enhancement projects

Interactions between enhancement projects and harvesting or other restoration actions in progress can have substantial impacts both on the performance and evaluation of enhancement projects. Restoration and other management activities should therefore be co-ordinated in such a way that interactions increase the universe of inference, rather than allowing unanticipated interactions to confound analysis. This may include basic 'common- sense' planning, such as ensuring that habitats of both the enhanced and control populations remain free from interference from logging or mining during the experimental evaluation period, and not starting enhancement projects at the same time as other restoration actions. For example, if culling of beavers or removal of beaver dams coincides with the introduction of incubation boxes, it will be difficult to separate out the respective individual effects of these restoration actions. For additional examples of how to use harvesting to create contrasting treatments, see section IIIB.

Large investments may hamper objective evaluation

The costs of enhancement infrastructure can act as a substantial roadblock to experimentation. These costs not only limit the types of experimental design that might be feasible and reduce the number of restoration alternatives that can be explored, but large investments required also tend to create social commitment and a feeling of entitlement that resist attempts at critical evaluation and potential corrective action. Thus, Wood (2002) argues that the fertilization projects undertaken as part of British Columbia's SEP were largely well-designed and executed and therefore successful in increasing our understanding of this type of enhancement because the costs involved were comparatively minimal. In contrast, the large investment required for building hatcheries created considerable pressure to produce quick results in terms of increased salmon output and limited political interest in experimentation. Given the costs and potential political complications involved in experimenting with capital-intensive enhancement projects, it seems advisable to rigorously assess potential benefits and weigh them against costs both in terms of research required to ascertain effects and in terms of risk to the natural population before embarking on such projects.

E. Socio-economic programs

Management of human attitudes and behaviours can be another important component of a broad salmon restoration program. Socio-economic initiatives such as education campaigns, co-management and community stewardship, economic incentives, or fisheries certification can play a useful role in restoration. However, such initiatives should be subject to the same standards of rigorous design and critical evaluation as projects focused on manipulating harvest rates, juvenile salmon abundance, or salmon habitat more directly. People do not normally think of socio-economic initiatives in the context of an experimental design, but they should because these are simply a different category of manipulation that is subject to unexpected outcomes.

As an example of a fairly typical socio-economic initiative, consider the Habitat Conservation and Stewardship Program (HCSP), which was launched by Canada's Department of Fisheries and Oceans in 1998 to enhance fish habitat stewardship in British Columbia. Considerable effort went into the design of this program, especially with respect to understanding the unique requirements of community organizations and community-based processes (Paish 1999). A study conducted in the initial phase of HCSP assessed existing knowledge, attitudes, and behaviour of BC citizens (Di Paula 2000). Formal program evaluations conducted in 2001 were based entirely on interviews with staff and program participants and primarily focused on the *perception* of participants as to whether the program was meeting its objectives (HCSP 2001a, 2001b). While these evaluations did provide some insights into whether the program was successful at building relationships with communities, there seems to have been no monitoring of habitat benefits or follow-ups to this initial survey that would allow direct inferences about the program's effectiveness in changing attitudes and behaviour of citizens, or about the program's effect on habitat and fish.

While a somewhat haphazard approach to program design and evaluation does not mean that the program is necessarily unsuccessful, application of the principles of experimental design should make it easier to identify the most efficient ways for delivering these programs and guide funding decisions over time. As with other types of restoration projects, this includes clearly stated, measurable objectives, rigorous monitoring and evaluation of the effect of a project against a control group and/or pre-treatment observations, and use of replication or other means to account for extraneous confounding factors. For example, a fairly basic BACI-style design for an education program that is intended to reduce by-catch of salmon in the trawl fishery might look as follows:

1. Select participating boats or operators, following the same principles one would use for selecting sampling plots in an ecological experiment.
2. Conduct a survey to assess participants' initial knowledge of salmon ecology, as well as knowledge of, and attitudes towards, salmon conservation concerns. If data on by-

catch are not already collected routinely for each vessel, initiate a monitoring program to track by-catch during the time of the study.

3. Divide participants into two groups, one of which will serve as a control, while the other will participate in a training course to learn about salmon conservation issues and options for reducing by-catch.
4. After completion of the training course, conduct another survey to re-assess knowledge and attitudes among the two groups. Contrast survey results and by-catch data obtained before and after the course to determine, whether the program was successful at changing attitudes and knowledge of the operators, and whether such changes in attitude in fact translated into changes in fishing behaviour and ultimately reduced by-catch. Comparison with results obtained for the control group accounts for various influences extraneous to the education program that might have affected participants' attitudes and behaviour.

Logical extensions of such an initiative to reduce by-catch might include additional treatment options, such as providing special incentives for one of the participating groups (e.g., a bonus of extra fishing time or public recognition for operators who consistently have low by-catch).

As with ecological restoration projects, conducting and evaluating socio-economic initiatives within an experimental design framework does not guarantee their success. However, it will help to discriminate approaches that work well from those that do not, and thus hopefully increase the effectiveness of restoration initiatives in the long run. Perhaps equally important, formal evaluation of socio-economic initiatives with respect to their ecological effects will put these approaches on a more equal footing with restoration actions that try to address conservation concerns through direct ecosystem manipulation, which in turn will allow more informed tradeoffs among these two approaches to restoration.

F. Tests of new technology or equipment

The backbone of salmon restoration projects is composed of technical know-how of personnel, including their use of engineered structures, scientific devices, and data collection protocols. For restoration projects to be successful, it is critical that these technical aspects of a project be thoroughly tested, verified, and their proper use documented. In some situations, testing and honing of skills and equipment may be incorporated as part of an applied restoration project. This has been the case in several river and lake fertilization experiments (e.g., McQueen 2002, Slaney et al. 2003, Wilson et al. 2003), where equipment and nutrient release protocols had to be changed or adjusted several times to achieve desired nutrient loadings. In cases such as these, when unproven technology is employed or familiar technology is used in novel settings, it might be advisable to explicitly plan and budget for a testing and calibration phase during which devices and protocols are tuned and technical glitches are worked out, because ecological responses during this period may be non-representative. The same cautions apply when projects are run with untrained personnel, as would typically be the case when local communities manage or assist in a restoration project that requires substantial knowledge of scientific methodology and/or specialized technology. Both project proponents and funding agencies should allow for an extended learning phase during project startup in these cases. Project implementation and monitoring schedules should be structured not to rely on ecological results during this initial phase, and should contain extra tests and consistency checks to uncover potential problems.

When testing and calibrating skill sets and technology or extrapolating from experiences gathered elsewhere, it is important to consider the full range of expected operating conditions and to assure that test conditions are representative of field conditions and/or conditions in previous applications. For example, failure to anticipate and account for the dynamic movement of real-world river channels in the design and placement of commonly used artificial habitat structures, such as log structures, spawning riffles, and bank stabilization, has often been identified as a key cause for the failure of such structures (Frissell and Nawa 1992, Kauffman et al. 1997, Roni et al. 2002). As with ecological responses to restoration projects, perhaps the most important lesson from past experience is that both a thorough initial assessment before installation *and* ongoing monitoring after installation are needed to maximize the chance of success and ensure prompt learning from failures.

IV. Strategic Overview Planning of Restoration Initiatives: Some Key Considerations for Managers and Funding Agencies Overseeing the AYK SSI Restoration Projects

The principles of good experimental design apply not only to individual restoration projects, but also to an overall restoration program to make it more successful. In essence, a restoration initiative may be viewed as a meta-experiment, in which individual restoration projects constitute individual treatments or treatment groups, and sometimes replicates. While opportunities for replication and controls may be limited, application of the other principles of experimental design that we have discussed are quite relevant at this broad program level, in particular:

- A clear statement of goals and objectives,
- Careful scoping,
- Complete measurement of responses (i.e., relevant indicators)

These principles will help the program administration to set a positive example of institutional culture and integrate individual projects into a cohesive program. Indicators can then be chosen to assist further with that integration across projects. An additional important element for administrators of the AYK SSI is to ensure extensive communication among various parties about the nature and benefit of experimental designs and methods to implement them. We summarize below some key ways in which good experimental design at all levels in program administration can contribute to the success of individual projects and the program as a whole.

Clear Objectives

A clear a priori statement of what constitutes success and how it should be measured is as critical for successful execution of a restoration initiative as it is for individual projects. In addition to formulating explicit restoration targets, restoration programs may benefit from setting explicit objectives to maximize the benefits from individual projects. One such additional objective should be to make the universe of inference from experimental results as wide as possible, by facilitating comparison between studies (meta-analysis), co-ordinated sampling protocols, and the widest possible use of available data (Marmorek et al. 2004; Paulsen and Fisher 2005).

Another additional objective should be to pursue an institutional culture of continuous learning and adaptation. Since the stakes tend to be higher for a whole restoration initiative than for individual restoration projects, explicit emphasis on learning from experience and avoiding mistakes of the past is even more important at this level. However, political or institutional pressure to produce results often prompts management agencies to create the appearance of success and "bury mistakes" (Hilborn and Walters 1981). Hilborn and Winton (1993) identified such unwillingness to learn as a major failing of the British Columbia Salmonid Enhancement Program in the 1970s and 1980s.

Their recommendations for avoiding this problem in the future are relevant to other large restoration-type programs, such as the AYK SSI. These recommendations include:

- external oversight to improve accountability,
- use of flexible technology to insure prompt discontinuation of projects or approaches that appear ineffective,
- allocation of substantial resources to monitoring and evaluation, and
- communication of realistic expectations.

In addition to these recommendations, making explicit program goals out of "continuous learning" and "flexible response to new information" could serve as a further incentive for maintaining accountability and creating an institutional climate that encourages good experimental design. Last, but not least, the program's objectives should be clearly communicated to project proponents and participants, and feedback needs to be channelled back promptly to program administrators to allow for effective monitoring and timely mid-course corrections, if necessary.

Appropriate Scope

Allocation of resources across projects and years is one of the most important responsibilities program coordinators face. One key challenge is to co-ordinate with proponents and regulatory agencies to distribute efforts in space and time in a way that minimizes undesired interactions and maximizes opportunities for inference and integration. This might include zoning approaches, such as establishment of "treatment-free" areas that may serve as controls for multiple projects (Routledge 2002), as well as co-ordinated timelines.

Another key challenge is to allocate funds to experimental design and monitoring in a way that reflects the need to manage uncertainty and risk. While it could be argued that a full experimental setup and thorough monitoring are always beneficial, doing this for every restoration project is not necessarily the most efficient way to allocate money and other resources. In general, projects require more investment in the experimental design phase if they:

- have a highly uncertain outcome or carry a high risk, such as projects with potentially widespread and/or long-lasting effects or potentially irreversible impacts, or
- are part of an adaptive management experiment.

On the other hand, projects require primarily implementation monitoring and otherwise do not need major investment in experimental design if they:

- are small in terms of spatial extent and/or potential range of effects, or
- are ‘routine’, in the sense that rely on proven technology and produce well-understood effects.

Complete measurement of responses and effective monitoring and analysis

Just as with setting program objectives, monitoring and analysis at the level of program delivery need to meet a dual purpose, to facilitate evaluation of individual projects and to allow effective evaluation of the program as a whole.

<u>Facilitating evaluation of individual projects and their synthesis</u>

There are several steps that can be taken to streamline evaluation of individual projects, while at the same time paving the way for integration of accumulated experiences through synthesis and formal meta-analysis. First, a good archive and ongoing collection of baseline data can avoid unnecessary sampling and in some cases may eliminate the need for separate pre-treatment observation, thus substantially cutting down on start-up time. Good baseline data are also essential for taking advantage of unexpected natural experiments, such as an oceanographic regime shift, since data from routine monitoring programs will often be the only pre-treatment data available in this case. Co-ordination of field sampling across projects and opportunistic gathering of data can further help to extend data coverage in an economical way. Such an approach is illustrated by the federal, state, and tribal fish and wildlife organizations in the Columbia River Basin, which have established the Collaborative Systemwide Monitoring and Evaluation Project (http://www.cbfwa.org/committees/csmep/default.cfm). The broad goal of this program is to improve co-ordination and integration of environmental and biological data relevant to salmonids, thereby leading to an improved framework for an answering critical monitoring and evaluation questions in the region (Parnell et al. 2005).

If data or indicators are to be shared between projects or used for meta-analysis, compatibility across restoration projects is essential. For instance, in an extensive review of river restoration projects across the U.S., Bernhardt et al. (2005) found that individual projects were inadequately tracked. Of the projects in their database, 20% had no listed goals, only 58% had records on project costs, and only 10% indicated that assessment or monitoring were occurring. Such shortcomings constrain opportunities for broader scale analyses and interpretations about restoration success and failures. Unfortunately, the emphasis on novelty in current scientific culture also prompts scientists to make studies as unique as possible and limits efforts to repeat what already has been done elsewhere. Bunnell and Huggard (1999), who examined this problem in the context of research on biodiversity in managed forests, list several examples of this point. One of them concerns indicators used in studies of habitat use by shrews: out of 120 indicators used in 11 separate studies, 76% of the indicators were unique to individual studies and only 8% were used by more than two studies! This lack of comparability among studies suggests that funding agencies are well advised to act proactively by taking explicit steps to promote sufficient overlap. This may be accomplished through standards and protocols for data collection and selection of a minimum set of indicators such as age structure and stock identification of adults, and abundance of fry or smolts, as well as adults. Other indicators can then be added as needed.

As an alternative, or in addition, the agency delivering a restoration program may also wish to set up its own standardized monitoring program and protocols (e.g. Collins 2003; Crawford 2004). This route has been taken by the Washington Salmon Recovery

Funding Board (SRFB) (http://www.iac.wa.gov/srfb/default.asp). The SRFB specifies monitoring standards for project proponents, as well as conducting its own monitoring program (SRFB 2003), an approach that should provide an information base ideally suited for meta-analysis, as well as seamless integration of objectives for individual projects and overall program objectives. However, the considerable initial effort required to work out indicators and data collection and analysis protocols that make sense for every project may limit the general applicability of such an approach, especially in situations where projects are very diverse. In those situations, a mix of standard monitoring elements and specialized design may be required to balance the need for compatibility with enough flexibility to accommodate the unique requirements of each project.

Last, but not least, data and results from individual studies must be documented, organized, and made accessible if they are to be useful. This requires adequate funding and proactive planning for effective archival and dissemination work. Similarly, synthesis and meta-analysis of the data collected and the insights gained in individual restoration projects should receive due consideration in budget allocation and Requests for Proposals (RFPs) to get the most out of the available information. For example, the SRFB allocates up to 20% of total project costs to monitoring. While there is no lower threshold to this figure, standard monitoring requirements and the independent monitoring executed by SRFB staff should presumably assure that all projects receive adequate monitoring coverage. SRFB also specifies clear criteria for reporting of monitoring results, including estimates of data quality and completeness, and pertinent information will be collected and made available through a common online database.

Program evaluation

Although there is typically overlap between objectives and results of a program and the objectives and results of the individual projects that it includes, there are (or should be) overarching program objectives that must be evaluated at the level of the program as a whole. As discussed previously, California's large restoration programs and British Columbia's Salmonid Enhancement Program (SEP) provide cautionary examples in this respect. Specifically, many of the projects initiated under SEP were carefully monitored and adequately evaluated (Hilborn and Winton 1993). However, evaluation of overarching SEP enhancement objectives was compromised because data on adult returns were not collected systematically, not always made available in a timely manner, and not always used during planning and evaluation (Hilborn and Winton 1993). As a consequence, its projects, though often well executed, produced neither an unequivocally successful restoration and enhancement program, nor a good learning/knowledge base documenting the capabilities and limitations of the various enhancement technologies. This example illustrates that, as with individual restoration projects, a clearly defined set of program objectives and indicators to quantify these objectives must be followed up with a monitoring program that allows not just for pro-forma evaluation, but for actual empirical measurement of the performance indicators. This may well require budgeting for collection of data and information above and beyond what is provided by individual projects.

V. Conclusion/Summary

A fundamental message of this report is that restoration of Pacific salmon populations in the Arctic-Yukon-Kuskokwim region should be conducted as part of carefully designed and co-ordinated programs that have learning as a clear goal, along with the primary goal of recovery of salmon populations. We cannot just take actions in the future that seem like they should help; we also have to monitor the results to determine whether they do help. In the context of restoration projects, there are three key benefits of applying the principles of experimental design:

- First, they will strengthen the conclusions drawn from attempts at restoration. Without applying such principles, there is a large chance of drawing incorrect conclusions due to confounding factors that changed while the restoration was under way. By using principles of experimental design, we are more likely to correctly determine in the shortest time which methods of restoring salmon populations work best.
- Second, an experimental design framework will use restoration budgets efficiently by reducing funding of ineffective actions over time and focusing resources on the more successful methods.
- Third, application of these principles will also contribute to a legacy of institutional and community-based knowledge about which restoration methods worked best and which did not, and in what situations.

Of course, proper experimental design, implementation, monitoring, and evaluation of restoration projects for salmon can be expensive. However, consider the alternative. There can also be enormous costs of not taking an experimental approach to these steps. Specifically, projects that are unsuccessful may be allowed to continue or even expand unchecked, using funds that might otherwise be available for more promising alternative projects. Instead, efforts should be shifted from ineffective projects to the most effective methods of restoration. An even worse outcome might arise from counterintuitive negative effects of some action that is allowed to continue, such as disease spread to wild stocks by supplementation with hatchery juveniles, which could severely reduce survival rate of the remaining wild population. Without a rigorous framework for evaluating restoration actions, such negative impacts will not be detected, let alone avoided in the future.

General recommendations

The following recommendations apply not only to manipulations of salmon habitat in fresh water but also to management regulations that affect harvesting, and socio-economic programs aimed at salmon conservation.

1. All salmon restoration actions should have at least two clearly stated main objectives:
   - Achieve results in terms of salmon abundance (specifying which populations, over what period, etc.). This objective could include either increasing salmon abundance or preventing loss of resilience or further reduction in abundance.
   - Learn over time which restoration actions work best and which do not.
2. All restoration projects affecting a particular population should be co-ordinated to reduce the chance of unintended interactions between projects. This also applies to other activities that may affect salmon directly or indirectly.
3. For the same reason as #2, co-ordinate actions *across* salmon populations and species.
4. Restoration actions should be carefully planned as part of a larger experimental design to reduce chances that naturally occurring confounding factors such as simultaneous changes in climate conditions will make it difficult to determine the effect of these actions. This can be achieved by using treatments and controls.
5. A carefully designed monitoring and evaluation program should be implemented to determine over the long term whether the restoration actions had the intended effects. This recommendation includes not only data collection but also data analysis and dissemination of results; the latter steps are often not funded sufficiently.
6. It is also critical to monitor how well the restoration actions were actually implemented.
7. Restoration programs should address habitat-forming processes. There is a critical distinction between restoration actions that merely "rehabilitate" site-specific conditions (such as spawning areas, pools, banks, large woody debris, boulders etc.) and those that aim to restore the natural processes that form, rehabilitate, and maintain salmon habitat over the long term. An ideal goal for watershed restoration would thus be to re-establish natural watershed processes (Ebersole et al. 1997). If met, this goal would also reduce the need for human intervention in the future.
8. Restoration programs should follow similar guidelines as individual restoration projects to remain accountable and promote an institutional culture where learning is encouraged. In particular, institutions should avoid falling into a "sunk cost" trap, where initially promising approaches continue to be pursued even if they prove ineffective because (a) initial investments were high and there is a political need to amortize these investments, or (b) individuals or organizations have developed a vested interest in the continuation of their projects.
9. Restoration projects that do not follow a complete experimental design are not necessarily unacceptable, but compromises should be avoided if they sacrifice too much power to discriminate among hypotheses and thereby lead to uninformative designs. While it is prudent to implement only the minimal design necessary for evaluation, skimping initially on good design might result in large costs later.

To conclude, we have argued strongly that despite the challenges of applying the principles of experimental design to restoration efforts in the Arctic-Yukon-Kuskokwim region, the benefits are substantial for doing so. Even if only some of the principles can be applied, there will be improvements in the ability to reliably interpret the effectiveness of various restoration actions on salmon. The more extensively that those principles can be implemented, the greater those improvements. The learning that will result will permit the AYK SSI funds to be used more effectively in the long term by focusing on those restoration activities (including fisheries management actions) that are most beneficial to the salmon populations. Application of experimental design principles now to the design of expanded baseline monitoring programs and scientific research that might be funded by AYK SSI will reap further benefits. By collecting data in the near future related to testing hypotheses about mechanisms causing change in AYK salmon abundance, people in the region will potentially be better able to respond relatively quickly to any future detrimental changes in that abundance.

APPENDIX: Checklist for Evaluating Proposed Experimental Designs

(1) Clearly-stated hypothesis or objective

- ☐ a. Objective meets AYK SSI strategic goals?
- ☐ b. Objective expressed in terms of some measurable indicators?
- ☐ c. Clear criteria given for meeting objective?
- ☐ d. Component objectives sufficient to address overall objective of project?

(2) Project scoping

- ☐ a. Selected study site appropriate for addressing objective?
  - ☐ i. Site is representative
  - ☐ ii. Site is a representative sub-sample of SSI area
  - ☐ iii. Site is an area of special concern
  - ☐ iv. Site is not fully representative, but offers other overriding benefits
  - ☐ v. Site serves as a ‘worst-case’ scenario for conservative test (e.g., for testing of new technology)
- ☐ b. Spatial and temporal scale of project appropriate?
  - ☐ i. Scale(s) of proposed modification/manipulation matches scale(s) of key ecological processes to be manipulated?
  - ☐ ii. Are intermediate or long-term climatic fluctuations and extreme events of potential concern? If yes, are such concerns addressed in ED and monitoring scheme?
- ☐ c. Existing scientific knowledge adequately summarized and incorporated into project design?
- ☐ d. Emphasis on learning rather than validation of current scientific knowledge?
  - ☐ i. High uncertainty about ecological processes and outcomes should be reflected in: high-power ED, designed to exclude as many confounding influences, as possible; extensive monitoring of potentially relevant environmental covariates as well as indicators of system response; representative study site(s)
  - ☐ ii. ‘Routine’ projects with well-established technology and ecological outcomes should be characterized by: emphasis on cost-effectiveness of implementation and monitoring scheme
- ☐ e. Key confounding influences identified and addressed in ED scheme?
  - ☐ i. Climatic trends and fluctuations a concern
    - ☐ 1. spatial replication
    - ☐ 2. staircase design
    - ☐ 3. inclusion of climate covariates
  - ☐ ii. Spatial environmental gradients a concern
    - ☐ 1. blocking of treatments
    - ☐ 2. inclusion of relevant environmental factors as covariates
  - ☐ iii. Substantial differences in environmental conditions across study area a concern
    - ☐ 1. stratification
    - ☐ 2. inclusion and explicit modeling of relevant environmental factors

- ☐ iv. Lack of independence between replicates (see (4) below)
- ☐ v. Are there human activities planned or in progress that might distort the outcome of the project? (e.g., potential interactions with other manipulative experiments?)
  - ☐ 1. accounted for in spatio-temporal allocation of treatments/controls
  - ☐ 2. control variables included to disentangle confounding influences

(3) Use of contrasting treatments and/or controls

- ☐ a. Control(s) and/or alternative treatment(s) included?
- ☐ b. Sites selected for treatments and controls sufficiently similar with respect to relevant characteristics and synchronized with respect to relevant processes (esp. critical if paired design is used)?

(4) Replication

- ☐ a. Spatial and/or temporal replication used?
- ☐ b. Replicates independent with respect to relevant characteristics?
  - ☐ i. If temporal replication: effect of potential temporal autocorrelation in samples considered and accounted for (e.g., adjust degrees of freedom in analysis, explicit inclusion of environmental covariates responsible for temporal correlation …)?
  - ☐ ii. If spatial replication: effect of spatial autocorrelation in samples considered and accounted for?
- ☐ c. Replicates sufficiently similar that additional variability introduced by using replication does not overwhelm benefits of including them?
- ☐ d. For a design based on contrasting treatments, do the replicates for each level of treatment meet these criteria?
- ☐ e. If stratification is used:
  - ☐ i. Selected stratification scheme appropriate, i.e., does it address key differences between sites that are expected to inflate variability in response?
  - ☐ ii. Reduction in variability large enough to warrant stratification?
  - ☐ iii. Number of samples in each stratum large enough to provide sufficient power?
- ☐ f. If staircase design or blocking is used:
  - ☐ i. Blocking scheme appropriate, i.e, does it aim to eliminate potential biases introduced by environmental trends/gradients?

(5) Randomization

- ☐ a. If randomization used, is sample size sufficient to avoid accidental biases?
- ☐ b. If partial randomization is used (e.g., random selection of sites within non-random blocks), is the proposed statistical analysis appropriate for the chosen sampling protocol?
- ☐ c. If no randomization used, are potential confounding variables adequately accounted for in the design?

(6) Sample size

- ☐ a. Is the sample size adequate to produce sufficiently precise estimates of parameters, including the response to the restoration action?
- ☐ b. Is the sample size adequate to produce sufficiently high statistical power, i.e. chance of detecting a statistically significant response of the magnitude that is deemed biologically or economically important?

(7) Complete measurement of responses

- ☐ a. Selected set of response variables/indices sufficient to test hypothesis and address study objective(s) by matching the scale of ecological processes being considered?
- ☐ b. Response variables selected to maximize ability to distinguish between confounding factors?
- ☐ c. Indices and/or measurement protocols follow standard guidelines?
- ☐ d. Spatial and temporal sampling protocol matches scales of expected variability in response variables and covariates? (i.e., sampling extent and frequency appropriate)
- ☐ e. If uncertain about a, b, d, conduct a pilot study or a simulation to resolve uncertainties?

(8) Effective monitoring and analysis

- ☐ a. A priori power analysis, pilot study, and/or simulation study to determine optimal sampling scheme?
- ☐ b. Adequate monitoring in place?
  - ☐ i. Effectiveness monitoring
  - ☐ ii. Implementation monitoring
- ☐ c. Contingency plans for likely implementation problems?
- ☐ d. Proposed analysis methods (statistical and otherwise) appropriate to data and sampling protocol?

Acknowledgments

We are grateful to the Arctic-Yukon-Kuskokwim Sustainable Salmon Initiative for the opportunity to conduct this review. We also thank several people who work in the AYK region and who provided valuable information, including Gene Sandone, Jim Minard, Karen Dunwall, Mark Gaboury, and Doug Molyneaux. In addition, we thanks Dave Marmorek, Ian Parnell, and Brent Mossop, who provided useful discussions.

References

AMFSTP (Adaptive Management Forum Scientific and Technical Panel). 2004. Final Report: Adaptive Management Forum for Large-scale Channel and Riverine Habitat Restoration Projects. Prepared for the United States Fish and Wildlife Service, Anadromous Fish Restoration Program and CALFED Bay-Delta Program. Available from: http://www.delta.dfg.ca.gov/afrp/project.asp?code=2000-45.

AYK SSI (Arctic-Yukon-Kuskokwim Sustainable Salmon Initiative). 2004. Protecting the future of salmon. 2004 Annual Report, 13 pp.

Bash, J.S. and C.M. Ryan. 2002. Stream restoration and enhancement projects: is anyone monitoring? *Environmental Management* 29(6): 877-885.

Bayley, P.B. 2002. A review of studies on responses of salmon and trout to habitat change, with potential for application in the Pacific Northwest. Washington State Independent Science Panel.

Beechie, T.J., G. Pess, E. Beamer, G. Lucchetti, and R.E. Bilby. 2003. Role of watershed assessments in recovery planning for salmon. Pages 194-225 in D.R. Montgomery, S. Bolton, D.B. Booth, and L. Wall (eds). Restoration of Puget Sound Rivers. University of Washington Press. Seattle, Washington.

Bernhardt, E.S., M.A. Palmer, J.D.Allan, G. Alexander, K. Barnas, S. Brooks, J. Carr, S. Clayton, C. Dahm, J. Follstad-Shah, D. Galat, S. Gloss, P. Goodwin, D. Hart, B. Hassett, R. Jenkinson, S.Katz, G.M.Kondolf, P.S. Lake, R. Lave, J.L.Meyer, T.K. O'Donnell, L. Pagano, B. Powell, and E. Sudduth. 2005. Synthesizing U.S. river restoration efforts. *Science* 308: 636-637.

Bisbal, G.A. 2001. Conceptual design of monitoring and evaluation plans for fish and wildlife in the Columbia River ecosystem. *Environmental Management* 28 (4): 433-453.

Bisson, P. A., R. E. Bilby, M. D. Bryant, C. A. Dolloff, G. B. Grette, R. A. House, M. L. Murphy, K V. Koski, and J. R. Sedell. 1987. Large woody debris in forested streams in the Pacific Northwest: past, present and future. Pages 143-190, *In:* E. O. Salo and T. Cundy (editors). Streamside Management: An Interdisciplinary Symposium on Forestry and Fishery Interactions. Fisheries Research Institute and College of Forest Resources, University of Washington, Seattle, Wash.,

Box, G.E.P., W.G. Hunter, and J.S. Hunter. 1978. Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building. John Wiley, New York, 653 pp.

Brannon, E.L., D.F. Amend, M.A. Cronin, J.E. Lannan, S. LaPatra, W.J. McNeil, R.E. Noble, C.E. Smith, A.J. Talbot, G.A. Wedemeyer, and H. Westers. 2004. The controversy about salmon hatcheries. *Fisheries* 29(9):12-31

Bunnell, F.L. and B.G. Dunsworth. 2004. Making adaptive management for biodiversity work - the example of Weyerhaeuser in coastal British Columbia. *The Forestry Chronicle* 80(1): 37-43

Bunnell, F. L., and D. J. Huggard. 1999. Biodiversity across spatial and temporal scales: problems and opportunities. Forest Ecology and Management 115:113-126.

Carpenter, S. R., T. M. Frost, D. Heisey, and T. K. Kratz. 1989. Randomized intervention analysis and the interpretation of whole-ecosystem experiments. *Ecology* 70:1142-1152.

Clemen, R.T. 1996. Making Hard Decisions: An Introduction to Decision Analysis. 2nd edition. Duxbury Press, Wadsworth Publishing Company. Belmont, California.

Collie, J.S., R.M. Peterman, and C.J. Walters. 1990. Experimental harvest policies for a mixed-stock fishery: Fraser River sockeye salmon, (*Oncorhynchus nerka*). *Canadian Journal of Fisheries and Aquatic Sciences* 47:145-155.

Collins, B.W. (editor). 2003. Interim Restoration Effectiveness and Validation Monitoring Protocols, California Coastal Salmonid Restoration Monitoring and Evaluation Program. California Department of Fish and Game, Fortuna, California. March 2003, 320 pp. Retrieved from: www.dfg.ca.gov/nafwb/pubs/ 2003/200303_Interim_Protocol_Manual.pdf.

Crawford, B.A. 2004. Protocol for monitoring effectiveness of in-stream habitat projects (Channel Reconfiguration, Deflectors, Log and Rock Control Weirs, Roughened Channels, and Woody Debris Removal). MC-2 Prepared for Washington Salmon Recovery Funding Board. Retrieved from: http://www.iac.wa.gov/Documents/SRFB/Monitoring/MC-2_Instream_Habitat_Projects.pdf

Dixon, W.J., and F.J. Massey. 1983. Introduction to Statistical Analysis. 4th ed., McGraw Hill Book Co., New York. 678 pp.

Di Paula, A. 2000. Benchmark Assessment of Public Awareness, Knowledge, Attitudes and Behaviour. Report prepared by Campbell Goodell Traynor Consultants for the Habitat Conservation and Stewardship Program. Published online at http://www-heb.pac.dfo-mpo.gc.ca/community/hcsp_e.htm.

Eberhardt, L.L., and J.M. Thomas. 1991. Designing environmental field studies. *Ecological Monographs* 61(1): 53-73.

Ebersole, J.L., W.J. Liss, and C.A. Frissell. 1997. Restoration of stream habitats in the western United States: Restoration as reexpression of habitat capacity. *Environmental Management* 21(1):1-14.

Eggers, D. M., L. R. Peltz, B. G. Bue, and T. M. Willette. 1991. Trends in abundance of hatchery and wild stocks of pink salmon in Cook Inlet, Prince William Sound, and Kodiak, Alaska. Professional Paper 35, Alaska Department of Fish and Game, Juneau, AK.

Fowler, C.W., and L. Hobbs. 2002. Limits to natural variation: implications for systemic management. *Animal Biodiversity and Conservation* 25(2): 7-45.

Frissell, C. A., and R. K. Nawa. 1992. Incidence and causes of physical failure of artificial habitat structures in streams of western Oregon and Washington. *North American Journal of Fisheries Management* 12:182-197.

Frissell, C.A., W.J. Liss, C.E. Warren, and M.D. Hurley. 1986. A hierarchical framework for stream habitat classification: Viewing streams in a watershed context. *Environmental Management* 10(2): 199-214.

Gardner, J., D. L. Peterson, A. Wood, and V. Maloney. 2004. Making Sense of the Debate about Hatchery Impacts: Interactions Between Enhanced and Wild Salmon on Canada's Pacific Coast. Prepared for the Pacific Fisheries Resource Conservation Council, Vancouver, B.C.

Green, R.H. 1979. Sampling Design and Statistical Methods for Environmental Biologists. John Wiley & Sons Ltd., New York. 157 pp.

Groot, C., and L. Margolis (editors). 1991. Pacific Salmon Life Histories. University of British Columbia Press, Vancouver, B.C.

Hairston, N.G. (ed). 1989. Ecological Experiments: Purpose, Design, and Execution. Cambridge Univ. Press, Cambridge, U.K. 370 pp.

Hammond, J.S., R.L. Keeney, and H. Raiffa. 1999. Smart Choices: A Practical Guide to Making Better Decisions. Harvard Business School Press. Boston, Ma.

Hard, J. J. 2000. Genetic interactions between wild and cultured salmon. Pages 8-18 in Gallaugher, P.C. Orr, editors. Aquaculture and the Protection of Wild Salmon, Workshop Proceedings. Continuing Studies in Science, Simon Fraser University, Burnaby, BC.

Hard, J. J. 2002. Genetic risks of hatchery salmon production to wild salmon. Pages 72-85 in Orr, C., P. Gallaugher, and J. Penikett, editors. Speaking for the Salmon: Hatcheries and the Protection of Wild Salmon. Simon Fraser University, Simon Fraser University, Burnaby, British Columbia.

Hartman, G.F., J.C. Scrivener, and M.J. Miles. 1996. Impacts of logging in Carnation Creek, a high-energy coastal stream in British Columbia, and their implication for restoring fish habitat. *Canadian Journal of Fisheries and Aquatic Sciences* 53 (Suppl. 1): 237–251.

HCSP (Habitat Conservation and Stewardship Program). 2001a. Habitat Conservation and Stewardship Program Field Evaluation. Published online at http://www-heb.pac.dfo-mpo.gc.ca/community/hcsp_e.htm.

HCSP (Habitat Conservation and Stewardship Program). 2001b. Building Bridges with First Nations. Published online at http://www-heb.pac.dfo-mpo.gc.ca/community/hcsp_e.htm.

Hilborn, R. 1992. Institutional learning and spawning channels for sockeye salmon (*Oncorhynchus nerka*). *Canadian Journal of Fisheries and Aquatic Sciences* 49:1126-1136.

Hilborn, R. 1999. Confessions of a reformed hatchery basher. Fisheries 24:30-31.

Hilborn, R., and D. Eggers. 2000. A review of the hatchery programs for pink salmon in Prince William Sound and Kodiak Island, Alaska. *Transactions of the American Fisheries Society* 129(2):333-350.

Hilborn, R., and D. Eggers. 2001. A review of the hatchery programs for pink salmon in Prince William Sound and Kodiak Island, Alaska: Response to comment. *Transactions of the American Fisheries Society* 130(4):720-724.

Hilborn, R., and C. J. Walters. 1981. Pitfalls of environmental baseline and process studies. *Environmental Impact Assessment Review* 2:265-278.

Hilborn, R., and J. Winton. 1993. Learning to enhance salmon production: Lessons from the Salmonid Enhancement Program. *Canadian Journal of Fisheries and Aquatic Sciences* 50:2043-2056.

Hillman, M., and G. Brierley. 2005. A critical review of catchment-scale stream rehabilitation programmes. *Progress in Physical Geography* 29:50-76.

Hobbs, N.T. 2003. Challenges and opportunities in integrating ecological knowledge across scales. *Ecology and Management* 181:223-238.

Holling, C.S., A. Bazykin, P. Bunnell, W.C. Clark, G.C. Gallopin, J. Gross, R. Hilborn, D.D. Jones, R.M. Peterman, J.E. Rabinovich, J.H. Steele, and C.J. Walters. 1978. Adaptive Environmental Assessment and Management. John Wiley & Sons, Chichester, United Kingdom, 377 pp.

Houlahan, J. 1998. Big problems, small science. *Conservation Ecology* [online] 2(1): R1. Available from: http://www.consecol.org/vol2/iss1/resp1/

Hurlbert, S. H. 1984. Pseudoreplication and the design of ecological field experiments. *Ecological Monographs* 54:187-211.

ISAB (Independent Scientific Advisory Board). 2002. Hatchery surpluses in the Pacific Northwest. *Fisheries* 27(12):16-27.

Kauffman, J. B., R. L. Beschta, N. Otting, and D. Lytjen. 1997. An ecological perspective of riparian and stream restoration in the western United States. *Fisheries* 22:12-24.

Keeley, E.R. and C.J. Walters. 1994. The British Columbia watershed restoration program: summary of the experimental design, monitoring, and restoration techniques workshop. Watershed Restoration Management Report Number 1, Ministry of Environment, Land and Parks and Ministry of Forests, Victoria, British Columbia. Retrieved from: http://srmwww.gov.bc.ca/frco/bookshop/docs/wrmr_1.pdf

Keeney, R.L. 1982. Decision analysis: An overview. *Operations Research* 30(5): 803-838.

Kershner, J.L. 1997. Chapter 8 - Monitoring and Adaptive Management. Pages 116-131, *In:* Williams, J.E., C.A. Wood, and M.P. Dombeck (editors). Watershed Restoration: Principles and Practices. American Fisheries Society, Bethesda, MD.

Landres, P.B., P. Morgan, and F. Swanson. 1999. Overview of the use of natural variability concepts in managing ecological systems. *Ecological Applications* 9(4): 1179-1188.

Larkin, P.A. 1972. A confidential memorandum on fisheries science. Pages 189-197, in B.J. Rothschild, editor. World Fisheries Policy: Multidisciplinary Views. Univ. Wash. Press. Seattle, Wash.

Leschine, T.M., B.E. Ferriss, K.P. Bell, K.K. Bartz, S. MacWilliams, M. Pico, and A.K. Bennett. 2003. Challenges and strategies for better use of scientific information in the management of coastal estuaries. *Estuaries* 26(4B): 1189-1204.

Levin, P. S., R. W. Zabel, and J. G. Williams. 2001. The road to extinction is paved with good intentions: negative association of fish hatcheries with threatened salmon. *Proceedings of the Royal Society of London*, *Series B*:1153-1158.

Lichatowich, J. 2002. Pacific Salmon beyond the crossroads: A strategy for the conservation of Pacific Salmon. Report prepared for the Gordon and Betty Moore Foundation, 32 pp.

MacDonald, L.H., A.W. Smart, and R.W. Wissmar. 1991. Monitoring guidelines to evaluate effects of forestry activities on streams in the Pacific Northwest and Alaska. U.S. Environmental Protection Agency, EPA 910/9-91-001, Seattle, Wash.

MacGregor, B.W., R.M. Peterman, B.J. Pyper and M.J. Bradford. 2002. A decision analysis framework for comparing experimental designs of projects to enhance Pacific salmon. *North American Journal of Fisheries Management* 22(2):509-527.

Marmorek, D.R., I.J. Parnell, M. Porter, C. Pinkham, C.A.D. Alexander, C.N. Peters, J. Hubble, C.M. Paulsen and T.R. Fisher. 2004. A Multiple Watershed Approach to Assessing the Effects of Habitat Restoration Actions on Anadromous and Resident Fish Populations. Prepared by ESSA Technologies Ltd., Vancouver, B.C. for Bonneville Power Administration, Portland, OR. 420 pp. Retrieved from: http://www.essa.com/watersh-rpt.pdf

Maser, C., and J. Sedell. 1994. From the Forest to the Sea: The Ecology of Wood in Streams, Rivers, Estuaries, and Oceans. St. Lucie Press, Delray Beach, Florida.

McAllister, M.K. and R.M. Peterman. 1992. Experimental design in the management of fisheries: a review. *North American Journal of Fisheries Management* 12:1-18.

McQueen, D. 2002. Nimpkish watershed sockeye enhancement using lake fertilization. Pages 8-13 in Robson, D., J. Penikett, and P. Gallaugher, editors. Proceedings from the Nutrients and Salmon Production Workshop, Centre for Coastal Studies, Simon Fraser University, Burnaby, BC.

Mellina, E., and S.G. Hinch. 1995. Overview of large-scale ecological experimental designs and recommendations for the British Columbia Watershed Restoration Program. Prepared for B.C. Ministry of Environment Lands and Parks and B.C. Ministry of Forests. WRP Project Report No. 1.

Morgan, P., G.H. Aplet, J.B. Haufler, H.C. Humphries, M.M. Moore, and W.D. Wilson. 1994. Historical range of variability: A useful tool for evaluating ecosystem change. *Journal of Sustainable Forestry* 2: 87-111.

Mossop, B. and M.J. Bradford. 2004. Importance of large woody debris for juvenile chinook salmon habitat in small boreal forest streams in the upper Yukon River basin, Canada. *Canadian Journal of Forest Research* 34(9): 1955-1966.

Myers, R.A., S.A. Levin, R. Lande, F.C. James, W.W. Murdoch, and R.T. Paine. 2004. Hatcheries and endangered salmon. *Science* 303:180.

NRC (National Research Council, United States). 2004. Developing a Research and Restoration Plan for Arctic-Yukon-Kuskokwim (Western Alaska) Salmon. The

National Academies Press, Washington, D.C., 165 pp. Pre-publication copy available on-line from www.aykssi.org/prod/index.htm.

Paish, Howard 1999. Getting Ahead of the Curve - Habitat Conservation and Stewardship: An Assessment of Community-based Processes and Organizations. Report prepared by Howard Paish & Associates for the Habitat and Enhancement Branch, Fisheries and Oceans Canada. Published online at http://www-heb.pac.dfo-mpo.gc.ca/community/hcsp_e.htm.

Parnell, I.J. 2002. Use of decision analysis to design a habitat restoration experiment. Master's thesis, School of Resource and Environmental Management, Simon Fraser University, Burnaby, B.C. 69 pp.

Parnell, I.J., M. Porter and D.R. Marmorek. 2005. Collaborative Systemwide Monitoring and Evaluation Project (CSMEP) – Year 1, Project No. 2003-036-00, Annual Report for FY 2004. Prepared by ESSA Technologies Ltd., Vancouver, B.C. on behalf of the Columbia Basin Fish and Wildlife Authority, Portland, OR. 13 pp. + appendices. Retreived from: http://www.cbfwa.org/Committees/CSMEP/Documents/Project/CSMEPFY04Annual%20Report013105.doc

Paulsen, C.M. and T.R. Fisher. 2005. Do habitat actions affect juvenile survival? An information-theoretic approach applied to endangered Snake River Chinook salmon. *Transactions of the American Fisheries Society* 134: 68–85.

Peterman, R.M. 1984. Density-dependent growth in early ocean life of sockeye salmon (*Oncorhynchus nerka*). *Canadian Journal of Fisheries and Aquatic Sciences* 41(12):1825-1829.

Peterman, R.M. 1987. Review of the components of recruitment of Pacific salmon. *In*: M. Dadswell, R. Klauda, C. Moffitt, R. Saunders, R. Rulifson, and J.E. Cooper (eds.), Common Strategies of Anadromous and Catadromous Fishes, American Fisheries Society Symposium 1:417-429.

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

Peterman, R.M. 1990b. The importance of reporting statistical power: the forest decline and acidic deposition example. *Ecology* 71(5):2024-2027.

Peterman, R.M. 1991. Density-dependent marine processes in North Pacific salmonids: Lessons for experimental design of large-scale manipulations of fish stocks. ICES Marine Science Symposium 192:69-77.

Peterman, R.M. and B. L. Antcliffe. 1993. A framework for designing aquatic monitoring programs. Pp. 227-251. *In*: A.P. Farrell (ed.), The Aquatic Resource Research Project: Towards Environmental Risk Assessment and Management of the Fraser River Basin. Technical Report for the British Columbia Ministry of Environment, Victoria, B.C., 451 pp.

Peterman, R.M. and M.K. McAllister. 1993. A brief overview of the experimental approach to reducing uncertainty in fisheries management--an extended abstract. *In*: S.J. Smith, J.J. Hunt, and D. Rivard (editors), Risk Evaluation and Biological Reference Points for Fisheries Management. *Canadian Special Publications of Fisheries and Aquatic Sciences* 120:419-422.

Peterman, R.M., B.J. Pyper, M.F. Lapointe, M.D. Adkison, and C.J. Walters. 1998. Patterns of covariation in survival rates of British Columbian and Alaskan sockeye salmon stocks. *Canadian Journal of Fisheries and Aquatic Sciences* 55:2503-2517.

Peterman, R.M., and J.L. Anderson. 1999. Decision analysis: A method for taking uncertainties into account in risk-based decision making. *Human and Ecological Risk Assessment* 5(2): 231-244.

Pyper, B.J., F.J. Mueter, R.M. Peterman, D.J. Blackbourn and C.C. Wood. 2001. Spatial covariation in survival rates of Northeast Pacific pink salmon (*Oncorhynchus gorbuscha*). *Canadian Journal of Fisheries and Aquatic Sciences* 58:1501-1515.

Pyper, B.J., F.J. Mueter, R.M. Peterman, D.J. Blackbourn and C.C. Wood. 2002. Spatial covariation in survival rates of Northeast Pacific chum salmon. *Transactions of the American Fisheries Society* 131(3):343-363 plus the erratum printed later in 131(4):796.

Reeves, G. H., D. B. Hohler, B. E. Hansen, F. H. Everst, J. R. Sedell, T. L. Hickman, and D. Shively. 1997. Fish habitat restoration in the Pacific Northwest: Fish Creek of Oregon. Pages 335-359 in Williams, J. E., C. A. Wood, and M. P. Dombeck, editors. Watershed Restoration: Principles and Practices. American Fisheries Society, Bethesda, Maryland.

Roni, P, and T.P. Quinn. 2001. Density and size of juvenile salmonids in response to placement of large woody debris in western Orgeon and Washington streams. *Canadian Journal of Fisheries and Aquatic Sciences* 58: 282-292.

Roni, P., T. J. Beechie, R. E. Bilby, F. E. Leonetti, and M. M. P. G.R. 2002. A review of stream restoration techniques and a hierarchical strategy for prioritizing restoration in Pacific Northwest watersheds. *North American Journal of Fisheries Management* 22:1-20.

Rose, K.A., and E.P. Smith. 1992. Experimental design: the neglected aspect of environmental monitoring. *Environmental Management* 16 (6): 691-700

Routledge, R. 2002. Conveners' Report. Pages 1-3 in Orr, C., P. Gallaugher, and J. Penikett, editors. Speaking for the Salmon: Hatcheries and the Protection of Wild Salmon. Simon Fraser University, Simon Fraser University, Burnaby, BC.

Ruggerone, G.T., M. Zimmermann, K.W. Myers, J.L. Nielsen, D.E. Rogers. 2003. Competition between Asian pink salmon and Alaskan sockeye salmon in the North Pacific Ocean. *Fisheries Oceanography* 12(3):209-219.

Sainsbury, K.J., R.A. Campbell, R. Lindholm, and A.W. Whitelaw. 1997. Experimental management of an Australian multispecies fishery: Examining the possibility of trawl-induced habitat modification. *In:* Global Trends: Fisheries Management, Pikitch, E.L., Huppert, D.D., Sissenwine, M.P. (eds.). American Fisheries Society Symposium 20:107-112.

Schmitt, R. J. and C. W. Osenberg (editors). 1996. Detecting ecological impacts: concepts and applications in coastal habitats. Academic Press, San Diego, California, 401 pp.

Schwarz, C.J. 1998. Studies of uncontrolled events. Pages 19-39 In: V. Sit and B. Taylor (editors), Statistical Methods for Adaptive Management Studies. Land Management Handbook No. 42. British Columbia Ministry of Forests, Research Branch, Victoria, British Columbia.

Sedell, J.R., P.A. Bisson, F.J. Swanson, and S.V. Gregory. 1988. What we know about large trees that fall into streams and rivers. U.S. Forest Service General Technical Report PNW-GTR-229:47-81.

Shields, F.D., C.M. Cooper_Jr, S.S. Knight, and M.T. Moore. 2003. Stream corridor restoration research: A long and winding road. Ecological Engineering 20: 441-454.

Slaney, P.A., and A.D. Martin. 1997. The Watershed Restoration Program of British Columbia: Accelerating natural recovery processes. *Water Quality Research Journal of Canada* 32(2): 325-346.

Slaney, P. A., B. R. Ward, and J. C. Wightman. 2003. Experimental nutrient addition to the Keogh River and application to the Salmon River in coastal British Columbia. Pages 111-126 in Stockner, J., editor. Nutrients in Salmonid Ecosystems: Sustaining Production and Biodiversity. American Fisheries Society, Bethesda, Maryland.

SRFB (Washington Salmon Recovery Funding Board). 2003. Monitoring and Evaluation Strategy for Habitat Restoration and Acquisition Projects (Draft Document). Washington Salmon Recovery Funding Board.

Stouder, D., P.A. Bisson, and R.J. Naiman (eds.). 1997. Pacific Salmon and Their Ecosystems. Chapman and Hall. New York. 685 pp.

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

Swanson, D.N. 1991. Chapter 5 – Natural Processes. Pages 139-179 in W.R. Meehan (ed). Influence of forest and rangeland management on salmonid fishes and their habitats. American Fisheries Society. Bethesda, Md.

Thomas, L. And C.J. Krebs. 1997. A review of statistical power analysis software. *The Bulletin of the Ecological Society of America* 78(2):126-138.

Underwood, A.J. 1991. Beyond BACI: experimental designs for detecting human environmental impacts on temporal variations in natural populations. *Australian Journal of Marine and Freshwater Research* 42:569-587.

Volpe, J. P. 2000. Do we know what we don't know? Atlantic salmon in British Columbia: A Review. Pages 28-33 in Gallaugher, P.C. Orr, editors. Aquaculture and the Protection of Wild Salmon, Workshop Proceedings. Continuing Studies in Science, Simon Fraser University, Burnaby, BC.

Walters, C.J. 1986. Adaptive Management of Renewable Resources. MacMillan. New York, NY.

Walters, C.J. 1997. Challenges in adaptive management of riparian and coastal ecosystems. Conservation Ecology [online] 1(2): Article 1. Available from: http://www.consecol.org/vol1/iss2/art1/

Walters, C. J., and R. Green. 1997. Valuation of experimental management options for ecological systems. *Journal of Wildlife Management* 61:987-1006.

Walters, C. J. and R. Hilborn. 1976. Adaptive control of fishing systems. *Journal of the Fisheries Research Board of Canada* 33:145-159.

Walters, C. J., J. S. Collie, and T. Webb. 1988. Experimental designs for estimating transient responses to management disturbances. *Canadian Journal of Fisheries and Aquatic Sciences* 45:530-538.

Weber, E. D., and K. D. Fausch. 2003. Interactions between hatchery and wild salmonids in streams: differences in biology and evidence for competition. *Canadian Journal of Fisheries and Aquatic Sciences* 60:1018-1036.

Wertheimer, A. C., and W. R. Heard. 2002. Effects of hatchery releases and environmental variations on wild-stock productivity: consequences for hatchery enhancement of pink salmon in Prince William Sound, Alaska. Pages 64-71 in Orr, C., P. Gallaugher, and J. Penikett, editors. Speaking for the Salmon: Hatcheries and the Protection of Wild Salmon. Simon Fraser University, Simon Fraser University, Burnaby, BC.

Wertheimer, A. C., W. W. Smoker, T. L. Joyce, and W. R. Heard. 2001. Comment: a review of the hatchery programs for pink salmon in Prince William Sound and Kodiak Island, Alaska. *Transactions of the American Fisheries Society* 130(4):712-720.

Whoriskey, F. 2000. Infectious salmon aneamia: A review and the lessons learned for wild salmon on Canada's east coast. Pages 46-51 in Gallaugher, P.C. Orr, editors. Aquaculture and the Protection of Wild Salmon. Simon Fraser University, Burnaby, B.C.

Williams, J.E., C.A. Wood, and M.P. Dombeck (editors). 1997. Watershed Restoration: Principles and Practices. American Fisheries Society, Bethesda, MD.

Wilson, G. A., K. I. Ashley, R. W. Land, and P. A. Slaney. 2003. Experimental enrichment of two oligotrophic rivers in south coastal British Columbia. Pages 149-162 in Stockner, J., editor. Nutrients in Salmonid Ecosystems: Sustaining Production and Biodiversity. American Fisheries Society, Bethesda, Maryland.

Wood, A. 2002. The history, goals and direction of the Salmonid Enhancement Program (SEP) in British Columbia. Pages 5-8 in Orr, C., P. Gallaugher, and J. Penikett, editors. Speaking for the Salmon: Hatcheries and the Protection of Wild Salmon. Simon Fraser University, Simon Fraser University, Burnaby, B.C.

Wright, Ronald. 2004. A Short History of Progress. House of Anansi Press, Toronto, Ontario.

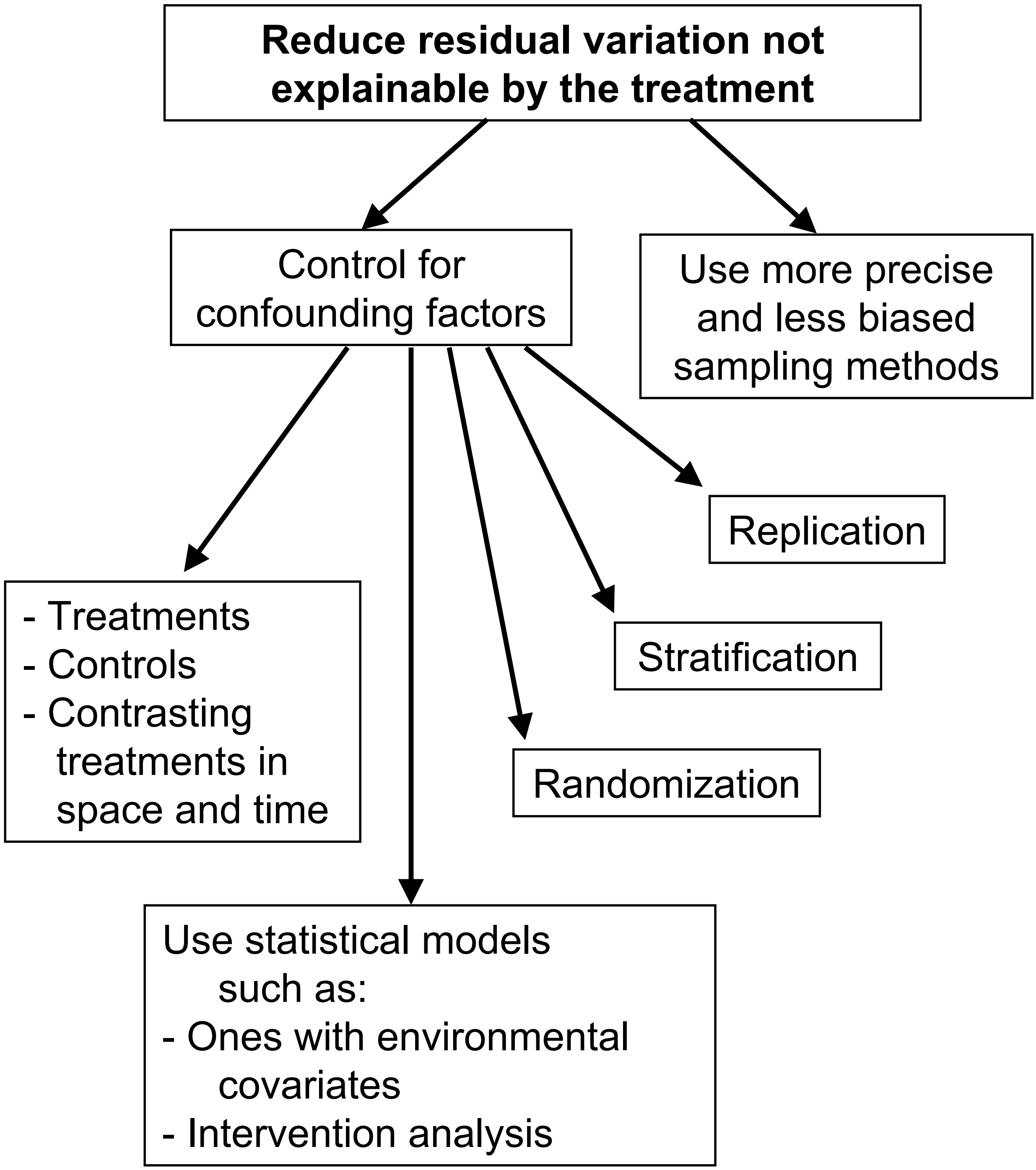

Figure 1. Methods that could be used to reduce the residual variation that is not explainable by the treatment imposed; in particular, reduce confounding factors.

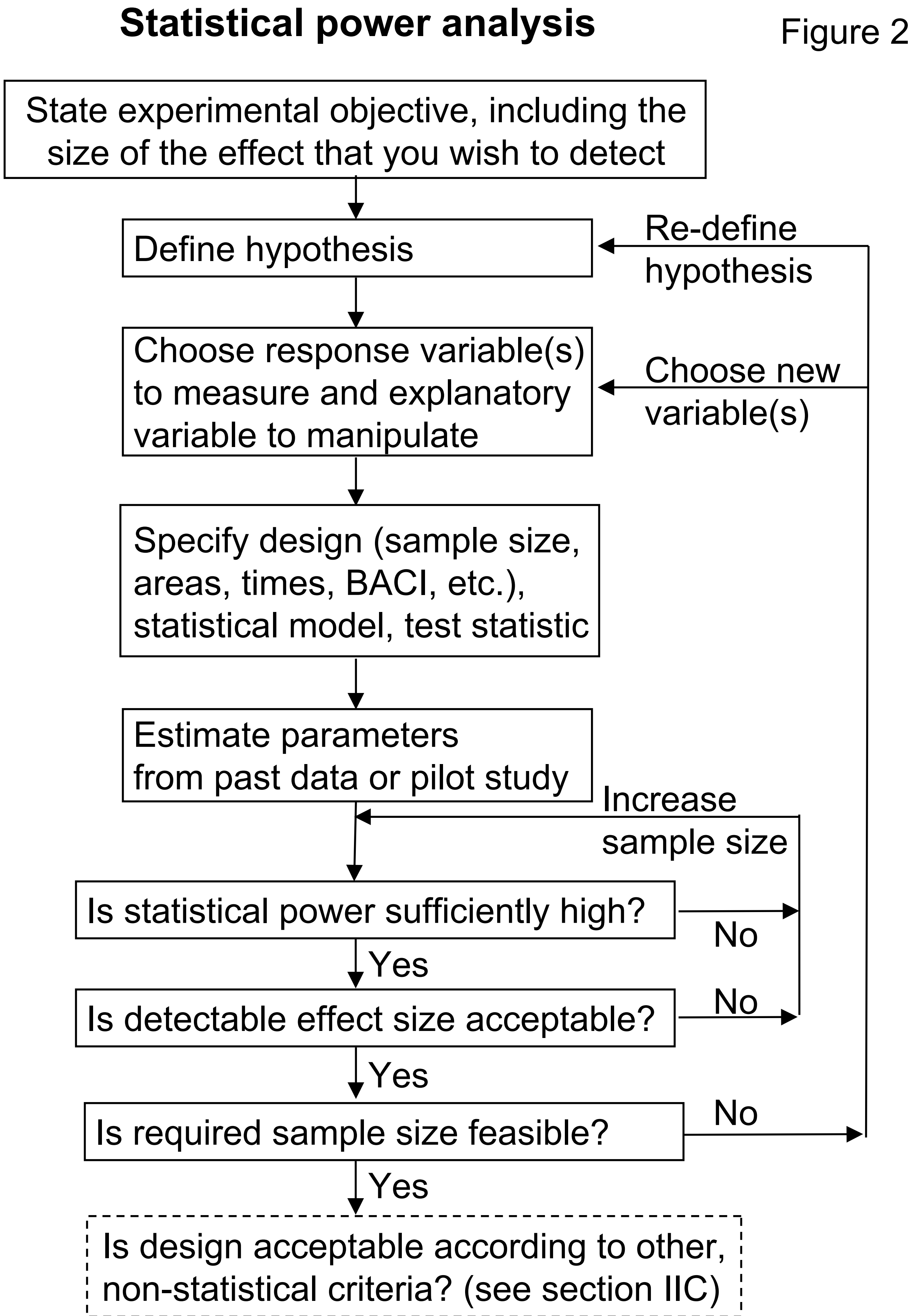

Figure 2. Steps in statistical power analysis.