COMPARATIVE SURVIVAL STUDY
OVERSIGHT COMMITTEE

Dr Nancy Huntley, Chairperson
Independent Scientific Advisory Board
1124 N. Stafford, APT. A
Arlington, Virginia. 22201

Dr. Eric Loudenslager, Chairperson
Independent Scientific Review Panel
Humboldt State University
Department of Fisheries
Arcata, California 95521

August 5, 2008

Dear Chairpersons Huntley and Loudenslager:

Since the inception of the state, federal and tribal fishery managers' collaborative Comparative Survival Study (CSS), both the ISAB and ISRP committees have reviewed CSS study design, analyses and reports. Over the past decade the CSS Oversight Committee has provided extensive presentations and analysis at the request of the ISAB and ISRP committees. The ISAB and ISRP reviews have helped improve the CSS study in numerous ways. As a matter of record, the CSS Oversight Committee has given the highest priority to addressing the comments and recommendations of the ISAB and ISRP committees. Consistent with our past approach to the ISRP/ ISAB comments and recommendations, the CSS Oversight Committee has given the recent comments and recommendations on the 10 year retrospective report careful consideration and has expended considerable time and effort researching the literature and body of scientific work in order to fully understand and consider the most recent comments. Our review is presented in the following discussion, which supports reconsideration of the ISRP/ISAB recommendation that resulted in the elimination of funding for tagging specific downstream mark groups from the CSS study. The CSS Oversight Committee believes that the elimination of tagging and analysis resulting from implementation of the ISAB/ISRP recommendation hinders scientific understanding of and resolution of the critical status of Chinook and steelhead populations in the Columbia Basin.

While the ISAB and ISRP Review (ISAB and ISRP 2007-6, hereafter "Review") of the CSS 10-year Retrospective Report (Schaller et al. 2007) was otherwise quite positive and helpful, the end result was the ISAB and ISRP recommendation that ongoing PIT-tagging efforts on Carson hatchery Chinook (an eleven-year time series of tagging efforts) should be terminated and that CSS-proposed PIT-tagging efforts on wild Chinook from the Warm Springs River (Deschutes River Basin) should not occur within the CSS. Following these recommendations, fisheries management agencies including the Washington Department of Fish and Wildlife (WDFW), the Oregon Department of Fish and Wildlife (ODFW), the Idaho Department of Fish and Game (IDFG), the Columbia River Inter-Tribal Fish Commission (CRITFC), and the United States Fish and Wildlife Service (USFWS) sent a letter (Joint Staff Letter, November 29, 2007) to the Northwest Power and Conservation Council staff (Jim Ruff) expressing their objections and concerns over the impacts and ramifications of these recommendations. The Council subsequently adopted the ISRP and ISAB recommendations by eliminating funding for these two tagging efforts in FY 2008 and 2009 from the CSS contract.

We find that the ISAB and ISRP recommendations to terminate existing and disallow proposed tag group monitoring data in the CSS disregard key management data needs that the CSS fulfilled and appear inconsistent with accepted scientific methodologies. In this response, we will show that:

- **The CSS study design is appropriate for measuring and monitoring life-cycle survival rates, for examining environmental factors associated with those rates, and for evaluating hypothesized mechanisms for variation in those rates,**
- **The measurement and inferential approaches used within the CSS are consistent with those used within an extensive body of peer-reviewed literature, including studies conducted by many of the ISAB/ISRP members themselves,**
- **The recommendations by the ISAB and ISRP to terminate existing (Carson hatchery Chinook) and to disallow proposed (Warm Springs wild Chinook) PIT-tagging efforts are inconsistent with their other Review recommendations, are inconsistent with past recommendations on other monitoring studies in the region, and are incompatible with several key management information needs,**
- **The scope of the CSS is primarily to establish a long-term dataset that measures the survival rate of annual generations of salmon and secondarily to examine empirical evidence relating to explanatory hypotheses for developing functional relationships useful for informing management decisions. In their Review, the ISAB and ISRP misinterpret the scope of the CSS as attempting to determine "unambiguous assignment of cause(s)" for all observed differences. These misinterpretations form the foundation for their rationale to eliminate the existing, and prevent the proposed, tagging efforts, and**
- **In addition to these issues associated with the elimination of monitoring tag groups, we would like to note several misconceptions apparent in the ISAB and ISRP interpretations of the CSS Report.**

The management issues affected by the ISAB and ISRP recommendations are simple, but critical. Upriver populations of spring/summer Chinook and steelhead populations have demonstrated poor life-cycle survival rates. Downriver populations of spring/summer Chinook salmon and steelhead have demonstrated much better, and in some cases self-sustaining, life-cycle survival rates. The upriver and downriver tagging efforts in the CSS are focused on determining where the life-cycle survival bottlenecks are occurring. Knowing where survival bottlenecks occur is critical for management decisions (Good et al. 2007). The ISAB and ISRP recommendations resulted in eliminating the possibility of the CSS fulfilling these critical management information needs.

We believe that several of these issues continue to reflect some basic misunderstandings about the CSS, and could have easily been resolved through discussion prior to decisions with long-term consequences being made. Unfortunately, the end result is that key, management-oriented information needs will not be met as a result of these ISAB and ISRP recommendations. Given the long-term implications and consequences of these recommendations, we believe that these recommendations warrant reconsideration. We are willing to discuss these issues with the ISAB and ISRP at any time.

The CSS study design, which includes downriver population groups and comparisons among population groups, is appropriate for measuring and monitoring life-cycle survival rates, for examining environmental factors associated with those rates, and for evaluating hypothesized mechanisms for variation in those rates.

First and foremost, the CSS is a management-oriented, large-scale monitoring study of spring/summer Chinook and steelhead. The foundational objective of the CSS is to establish a long-term dataset that measures the survival rate of annual generations of salmon from their outmigration as smolts to their return to freshwater as adults to spawn (smolt-to-adult return rate; SAR) (Schaller et al. 2007, p. 1). Through PIT-tags, survival rates and other important demographic responses (e.g., migration rates, migration timing) can be partitioned over the smolt-to-adult life-cycle (e.g., during the juvenile freshwater migration, during ocean residency, and during adult upriver migration). PIT-tags provide much finer level of resolution of monitoring information than can be obtained by less intensive methods (e.g., stock-recruitment data). By design, the CSS PIT-tagging efforts address several of the basin-wide monitoring needs, providing the "basic data on performance of upriver and downriver stocks [that] remain of value in monitoring and evaluation" (ISAB and ISRP 2007-6). Monitoring survival rates over the life-cycle can help identify where survival bottlenecks are occurring, which are critical for informing management decisions (Good et al. 2007). Again, the primary objective of the CSS is measurement of survival rates and other demographic responses for upriver and downriver stocks of Chinook salmon and steelhead in the Columbia Basin.

The CSS is consistent with a number of study design frameworks. Under the classification of Hurlbert (1984), the CSS is a comparative mensurative experiment. Under the classification of Rasmussen et al. (2001), the CSS could be considered an example of a time series intervention analysis. Under the classification of Roni et al. (2005), the CSS shares characteristics of both the extensive post-treatment design (through the monitoring of multiple population groups, with population groups providing spatial replication) and the intensive post-treatment design (through the monitoring of population groups over multiple years, with population groups providing both spatial and temporal replication). Under the classification of ISRP (2005-14), the CSS can be considered an effectiveness monitoring study, where the objective is to evaluate the effectiveness of various management actions (e.g., mainstem operations, mainstem and ocean environmental conditions, and transportation). However, the CSS design and analytical methods are most consistent with that of an "observational study" (Cochran 1983, Eberhardt and Thomas 1991, McDonald et al. 2007). Cochran (1983, p. 1) defines an observational study as having two characteristics:

> "1) The objective is to study the causal effects of certain agents, procedures, treatments, or programs.
> 2) For one reason or another, the investigator cannot use controlled experimentation, that is, the investigator cannot impose on a subject, or withhold from the subject, a procedure or treatment whose effects he desires to discover, or cannot assign subjects at random to different procedures."

Observational studies are used extensively within the fields of ecology, toxicology, paleontology, geology, and epidemiology in particular (Cochran 1983, Eberhardt and Thomas 1991, Rothman and Greenland 1998, Woodward 2005, Jewel 2005). Cochran (1983, p. 2) describes that "a basic difference between observational studies and controlled experiments is that the groups of people whom the investigator wishes to compare are already selected by some means not chosen by the investigator." Examples of observational studies include epidemiological studies that measured and compared lung cancer rates among population groups of smoking and non-smoking individuals. Through those comparisons, the effects of smoking and other risk factors could be examined, despite the ethical and logistic difficulties of randomly assigning individuals to the smoking or non-smoking treatment groups (Eberhardt and Thomas 1991). Analogously, it is obvious that population groups monitored within the CSS cannot be randomly assigned to different hydrosystem entrance locations or assigned to pre-determined environmental conditions (e.g., various levels of flow, spill, ocean upwelling conditions). The observational study design simply measures the biological responses of population groups to the conditions (i.e., treatment effects or management actions) that they experience. Thus, the CSS monitors the differences and similarities among population groups caused by the "natural experiment" within the Columbia Basin (Diamond 1986). However, the inability to randomly assign population groups to treatments or management actions does not preclude valid measurements of biological responses for those population groups or valid comparisons of biological responses among groups (Hurlbert 1984, Jewel 2005).

Diamond (1986) defines a "natural experiment" as a study where the "experimenter does not establish the perturbation but instead selects sites where the perturbation is already running or has run." In the selection of population groups to monitor, "along with the experimental sites, the investigator selects control sites so that the two types of sites differ in the presence or absence of the perturbation but are as similar as possible in other respects" (Diamond 1986). This selection of groups for study that are as similar as possible is known as "matching" or "blocking," and is one of three techniques for reducing the potential for confounding in observational studies (Cochran 1983). Examples of matching and blocking within the CSS include comparisons between population groups of the same species (matching species groups), of the same rearing type (matching rearing type- hatchery or wild), and of the same arrival timing at Bonneville Dam (temporal blocking). The other two techniques described by Cochran (1983) for reducing the potential for confounding are refinements in technique to reduce measurement error and control during the statistical analysis. Both of these techniques are applied in the CSS. Examples of efforts to reduce measurement error include the use of PIT-tag data, which have high precision relative to other measurement methods, and statistical analyses to remove the effects of sampling error from SARs (Chapter 4, Schaller et al. 2007). Examples of control during the statistical analysis include accounting for transportation location in the analysis of adult upriver migration success (Chapter 6) and analyses of the fork length, tributary emigration timing, and first-to-third dam migration rate relative to water travel time (Chapter 5).

Many of those environmental conditions that may influence the observed biological responses are measurable, varying across and within years, providing contrast in the environmental conditions experienced. If time series data of sufficient length are collected over contrasting environmental conditions, the development of functional correlations between the measured treatment effects and candidate associative factors may be possible (ISAB 2003-1, ISAB 2006-3, ISRP 2007-8). Statistical adjustment techniques such as regression and analysis of covariance provide analytical tools that can account for, and potentially remove, the effects of confounding factors in observational studies (Cochran 1983, Rothman and Greenland 1998, Woodward 2005, Jewel 2005).

The CSS provides extensive analyses that examine multiple hypotheses that attempt to explain the observed patterns of differential mortality between upriver and downriver stocks as well as hypotheses on the effects of freshwater and marine habitat conditions on the survival and migration rates of population groups (Chapter 5, Schaller et al. 2007). These evaluations are conducted to determine functional relationships that may be useful for informing management decisions (ISAB 2003-1, ISAB 2006-3, ISRP 2007-8). The CSS maintains "multiple working hypotheses" (Chamberlin 1890) for investigating these differences among population groups, migration experience, and over time. Our intent is to determine the support that the data offer for each competing model or hypothesis (Hilborn and Mangel 1997, Burnham and Anderson 2002). We agree with and apply the view of Hilborn and Mangel (1997) that "science consists of confronting different descriptions of how the world works with data, using the data to arbitrate between the different descriptions, and using the "best" description to make additional predictions or decisions."

While investigating candidate hypotheses and examining functional relationships are objectives of the CSS, we recognize that the identification of functional correlations does not establish causality. Rothman and Greenland (1998) provide an excellent discussion of the issues surrounding causality in scientific investigations. They state that "perhaps the most important common thread that emerges from the debated philosophies is Hume's legacy that proof is impossible in empiric science" and that "even when they are possible, experiments (including randomized trials) do not provide anything approaching proof." As a result,

> "All of the fruits of scientific work, in epidemiology or other disciplines, are at best only tentative formulations of a description of nature, even when the work itself is carried out without mistakes. The tentativeness of our knowledge does not prevent practical applications, but it should keep us skeptical and critical, not only of everyone else's work but of our own as well."

They contend that the appropriate scientific response to the tentative nature of these formulations is to “focus on testing the negation of the causal hypothesis” and that “observations can provide crucial tests of competing non-null causal hypotheses.” The methods and analyses used within the CSS are consistent with these views. None of the analyses in the CSS (or any other study in the region, including fully randomized experiments) are capable of proving cause-effect relationships (Rothman and Greenland 1998). However, consistent with accepted analytical techniques, extensive efforts were undertaken in the CSS to evaluate, examine, and in some cases negate proposed causal hypotheses for observed differences between population groups. The CSS examinations of the evidence for various hypotheses using assembly rules constitutes a “weight-of-evidence” approach, which has been recommended and applied to examine causal links for treatment effects in observational studies (Fox 1991, Gilbertson 1997, Beyers 1998, Lowell et al. 2000, Morales et al. 2003, Adams 2003, Brown et al. 2003, Collier 2003, Culp and Baird 2006).

Several of the key hypotheses relating to upriver/downriver comparisons that were examined in the CSS Report included investigations of differences in smolt emigration timing, differences in smolt size, differences in migration rate, and differences in smolt arrival timing to below Bonneville Dam. These were the primary hypotheses that had been proposed for explaining the observed differences between upstream/downstream stocks. Despite finding little evidence in support of these primary hypotheses, the ISAB and ISRP speculate that additional hypotheses are the cause of the observed differences and conclude, without scientific analysis, that it would be impossible to examine the degree of evidence for those hypotheses. The hypotheses proposed by the ISAB and ISRP (many of which could be addressed in future analyses) include differences in productivity, predator populations, local climatic conditions, life history, and “that the salmon have likely adapted and evolved (as well as they can) in the time elapsed” and that “Snake River stocks are, in all likelihood, no longer the same as their pre-hydrosystem ancestors.” However, recent genetic evidence from the standardized region-wide database indicates that John Day and Snake Basin spring/summer Chinook stocks are closely related (Seeb et al. 2007, Fig. 3). This study provides evidence that Snake Basin population groups have not diverged from the John Day population groups, which migrate through fewer dams.

Making comparisons among population groups and the inferential approaches used within the CSS are consistent with those used within an extensive body of peer-reviewed literature, including studies conducted by many of the ISAB/ISRP members themselves.

Measuring and making comparisons in survival, productivity, and other demographic measures among salmonid population groups is a common, well-established, and accepted scientific methodology for drawing inferences on factors associated with measured variation, <u>both within the Columbia Basin</u> (Schaller et al. 1999; Botsford and Paulsen 2000; Deriso et al. 2001; Petrosky et al. 2001; Budy et al. 2002; Wilson 2003; Zabel and Achord 2004; McHugh et al. 2004; Scheuerell 2005; Schaller and Petrosky 2007) <u>and across the northeast Pacific Ocean</u> (Adkison et al. 1996; Botsford and Lawrence 2002; Bradford 1995; Brodeur et al. 2004, 2007; Fisher et al. 2007; Hare et al. 1999; Hodgson et al. 2006; Holt and Peterman 2004; Mantua et al. 1997; Morris et al.

2007; Mueter et al. 2002a, 2002b, 2005, 2007; Orsi et al. 2007; Peterman et al. 1998, 2003; Peterman 2004; Pyper et al. 1999, 2001, 2002, 2005; Pyper and Peterman 1999; Schindler et al. 2003; Su et al. 2004; Trudel et al. 2007a, 2007b; Quinn et al. 2005). The spatial scales of these comparisons range from small (< 10 km) to large (up to 3,000 km) in the distances between population group spawning locations. Within the Columbia Basin, the range in distances and level of complexity between population groups for comparisons are obviously much smaller than those that have been examined across the northeast Pacific Ocean and published in the peer-reviewed literature.

Consistent with the scientific methods used within this body of work on salmonid population group comparisons, for over fifteen years researchers have compared and identified differences in the survival and productivity of upriver and downriver spring/summer Chinook stocks in the Columbia Basin using stock-recruitment data (Petrosky and Schaller 1992, 1996; Schaller et al. 1996, 1999; Deriso et al. 1996, 2001; Botsford and Paulsen 2000; Schaller and Petrosky 2007). Obviously, these studies have met the scientific standards of the peer-review process and were published in the peer-reviewed literature, in opposition with the viewpoints of the ISAB and ISRP that making comparisons among these same population groups is compromised by "inevitable confounding of all differences" (ISAB and ISRP 2007-6) and therefore should not occur within the CSS. Additionally, the primary criticism against using stock-recruitment data is that the data are imprecise and require assumptions about the stock-recruitment relationship. However, the CSS overcomes these issues by directly estimating differential survival through PIT-tag SARs. That is, no stock-recruitment relationship is necessary and the PIT-tag survival estimates are highly precise.

Consistent with the peer-reviewed literature cited above, the CSS measures and compares survival rates within and among salmonid population groups and across migration experiences. Contrary to the perspectives of the ISAB and ISRP, the process of making comparisons does not require that all potential causative factors must be controlled and accounted for (Hurlbert 1984, Jewel 2005). It is simply not possible to quantify or eliminate all possible interacting factors affecting species in a large ecosystem. However, the basic process of scientific inquiry is to use observational data to examine the degree of evidence for or against proposed hypotheses on the putative causes of observed variation. The ISAB and ISRP (2007-6) state that comparisons between population groups in the CSS do not meet scientific review criteria due to the inability to "unambiguously assign cause(s)" (p. 9), populations are not replicates of other populations (p. 12), "the inevitable confounding of all differences" (p. 12), and "the system is too complex" (p. 9). However all 39 of the publications listed above used similar observational study approaches as the CSS and met the scientific criteria of highly respected journals. Other comparative observational studies conducted by ISAB and ISRP members themselves using observational data (e.g., Poe et al. 1991; Bilby et al. 1998, 2003; Pearcy et al. 1996; Sork et al. 2005; Weeder et al. 2005) also applied observational study approaches and met the scientific criteria of peer-reviewed journals.

It appears that the ISAB and ISRP apply these criteria and judge them to be insurmountable flaws of one minor aspect of the CSS objectives. These criteria are then used as justification for their recommendations to terminate and disallow the collection of key monitoring data that are needed for the identification of survival bottlenecks and the evaluation of candidate hypotheses. The recommendation that "the additional monitoring and evaluation of downstream stocks should not be directed toward upstream-downstream comparisons" stands in stark contrast against the vast body of peer-reviewed and published scientific work on salmonid population dynamics and the comparisons made therin.

The recommendations by the ISAB and ISRP to terminate existing (Carson hatchery Chinook) and to disallow proposed (Warm Springs wild Chinook) PIT-tagging efforts are inconsistent with their other Review recommendations, are inconsistent with past recommendations on other monitoring studies in the region, and are incompatible with several key management information needs.

As stated above, the primary objective of the CSS is to establish a long-term dataset that measures the survival rate of annual generations of salmon from their outmigration as smolts to their return to freshwater as adults to spawn (smolt-to-adult return rate; SAR). Objectives 1-3 of the CSS proposal, which met "scientific review criteria," describe how that long-term dataset is to be established using each tagging group proposed, including the downriver tagging groups. At points in the Review, the ISAB and ISRP appear to recognize the value of these downriver monitoring data collected by the CSS. For example, they note:

> "future effort on downriver stocks might be directed at monitoring and evaluation of wild stocks, including determination of SARs and other population metrics for these stocks. This would be a logical part of a regional monitoring and evaluation program to inform managers of stock status and could be useful for assessing recovery of threatened and endangered stocks."

Because these first three CSS objectives apply to the downriver tagging groups proposed, and these three monitoring objectives "meet scientific review criteria," then therefore the downriver tagging groups must meet scientific review criteria on the merits of their monitoring value alone. The ISAB and ISRP recommendations to terminate the existing and disallow the proposed downriver tag groups in the CSS is therefore inconsistent with their support of the CSS monitoring objectives 1-3 in the Review.

Based on previous ISRP reviews, the region has invested in a long-term (11-year) time series of monitoring data for the Carson hatchery stock. Terminating this time series and disallowing the collection of other lower-river tagging time series within the CSS precludes the identification of where life-cycle survival bottlenecks are occurring, and knowing where survival bottlenecks occur is critical for management decisions (Good et al. 2007). Terminating and disallowing the lower river tagging efforts also precludes future investigations on the factors that may be affecting the performance of these stocks. These recommendations directly conflict with the guidance provided by the ISAB themselves on using multiple population groups to examine causal links (McDonald et al. 2007):

> "Monitoring multiple, independent, but similar observational studies, and analyzing them as if they were replicates subject to study-specific random effects, may provide convincing evidence for the effect of an action if results are consistent across the studies. In an inductive sense, monitoring numerous observational studies can suggest causal relations (e.g., Shipley 2000)."

It appears that the ISAB and ISRP have been inconsistent in their criticism of comparisons between population groups in the CSS relative to their recommendations on other large-scale monitoring efforts. As mentioned above, the CSS is an example of a large-scale effectiveness monitoring study and is classified as a comparative mensurative experiment. As defined by the ISRP (2005-14; page 20), the objective of an effectiveness monitoring study is:

> "Establishment of mechanistic or causal links between management actions and population responses with conclusions justified by replicated results and subjective judgment."

The ISRP (2005-14; page 25-26) then go on to state that when mensurative experiments are replicated over space and time,

> "corroborative results of the studies can provide compelling evidence for general conclusions. Such a mensurative study can be quite powerful and quite convincing when many replications of experimental results occur, i.e. there is establishment of the same relationships over several or many smaller studies. In this inductive sense, census and statistical monitoring in mensurative experiments do support research on cause of effects. The ISRP believes that this is the best study design for establishment of effectiveness of management actions in a large ecosystem such as the Columbia River basin. However, such conclusions require subjective judgment."

The two main points of the ISRP appear to be that: 1) powerful, convincing and compelling evidence for causal links can be established when mensurative experiments are replicated over space and time, and 2) some degree of subjective judgment must be used in interpreting resulting evidence on potential causal links arising from effectiveness monitoring studies. With these points, we fully concur. However, when considering the mensurative experiments replicated over space and time within the CSS, the ISRP and ISAB (2007-6) conclude, without scientific analysis, that comparisons within the CSS insurmountably preclude

> "unambiguous assignment of cause(s) impossible even if convincing, statistically significant differences in fish performance were established between upriver and downriver stocks. In sum, the system is too complex, and the possible sampling design necessarily too constrained in time and place, to reach conclusive findings on causation from this type of comparison."

These polar viewpoints are highly inconsistent. Either mensurative experiments replicated over space and time, combined with subjective judgment, are capable of providing compelling evidence of causal links or they are not. If they are, then the CSS mensurative experiments replicated over space and time, combined with subjective judgment, are capable of examining evidence for causal links. If they are not, then the recommendations provided by ISRP (2005-14) and McDonald et al. (2007) are false. Instead of supporting increased spatial and temporal replication efforts, as recommended by ISRP (2005-14), ISAB (2006-3), ISRP (2006-6), and McDonald et al. (2007) through the Carson and Warm Springs tagging efforts, the ISAB and ISRP recommend the termination or prevention of those replication efforts. Additionally, we found little or no evidence that the ISAB or ISRP found the scientific judgment used in the CSS was in any way inappropriate. In ISAB and ISRP (2007-6) there was no criticism of the extensive analyses that examined the main hypotheses on differences between upriver and downriver smolt characteristics, smolt responses to environmental conditions, or the temporal blocking comparisons between groups arriving at Bonneville at the same time. There was no criticism that the scientific judgment used by the CSS authors overstepped proper scientific interpretation or was in any way inappropriate when interpreting the results from these comparisons. But there also was no allowance for continued investigation of the newly-postulated hypotheses offered by the ISAB and ISRP to explain away the differences between upriver and downriver population performance. Instead, the ISAB and ISRP recommend terminating the future temporal replication that would have been provided by the Carson Hatchery tagging efforts (an 11-year time series) from the CSS and did not support augmenting the spatial and temporal replication efforts that would be provided by the Warm Springs wild Chinook tagging proposed by the CSS. These recommendations against replicating mensurative experiments over space and time directly oppose what the ISRP (2005-14) believes to be the "best study design for establishment of effectiveness of management actions in a large ecosystem."

We are aware of several other examples of the ISAB and ISRP supporting the use of comparative reference groups for making inferences using observational studies. The ISRP has supported the methodologies of project 200311400, a study which makes upriver/downriver (Snake vs. Yakima) comparisons between hatchery spring Chinook survival and migration rates using small numbers of radio tags (ISRP 2006-6). In ISRP (2004-8), the ISRP compliment a study design that uses CWT fish from a downriver hatchery stock released into the estuary (Blind Slough) to improve understanding of the effects of the ocean environment on upriver (Snake) spring Chinook stocks (study title: "Evaluation of the relationship among time of ocean entry, physical, and biological characteristics of the estuary and plume environment and adult return rates"). Another example is the ISRP (2007-1) review of project 200001400 where they support the downriver monitoring of lamprey populations, because this monitoring would be "providing the opportunity to compare trends in abundance of lamprey populations not affected by mainstem dams with those occurring above the dams." In ISAB and ISRP (2005-15), evaluating the effects of supplementation "involves contrasting trends in treatment and reference locations" even though they recognize that "treatment and reference locations will undoubtedly differ from each other beyond the supplementation treatment." Obviously, reference streams are not perfect replicates of supplementation streams, but despite this limitation the ISAB and ISRP recommend that researchers compare and contrast trends between reference and supplementation streams. In contrast, when the ISAB and ISRP consider the CSS comparisons, their "core reason a contrast of salmon survival between upriver and downriver locations is not advised is that the populations in tributaries downriver of the dams are not replicates of the upper Snake River populations" (ISAB and ISRP 2007-6; page 12). How is this criticism of the CSS comparisons consistent with the ISRP's support of Snake/Yakima comparisons in ISRP (2006-6) or the ISAB and ISRP support of supplementation versus reference stream comparisons in ISAB and ISRP (2005-15)? All are examples of observational study comparisons between sites or groups that are not perfect replicates of each other in all other factors beyond the treatment effect.

However, the most dramatic reversal in opinion on the utility of reference groups came from ISRP (2006-6) themselves in their review of the CSS proposal. In regard to the downriver tag groups, the ISRP (2006-6) said the following (emphasis added):

> "For this upriver-downriver comparison to be generally accepted, it seems prudent to **add more downriver sites in the future**. In response, **the CSS will add another downriver site in the Warms Springs River for wild Chinook tagging for 2007** to complement the ongoing tagging in the John Day River. **This is a positive action, however, additional downriver hatchery sites are even more important to add** because at this time, five upriver hatcheries are being used as tagging sites and only one downriver. There needs to be better hatchery to hatchery comparisons, and **adding several lower river hatcheries which show a range in return rates will provide a more realistic comparison in survival rates**."

At that time, the ISRP appeared to recognize that additional downriver tag groups would greatly improve understanding of between-population variation among the downriver population groups and between upriver and downriver groups. This improved understanding would also help identify life-stages of poor survival among populations (Good et al. 2007). In terms of additional downriver hatchery sites, Schaller et al. (2007; page 176) responded to the ISRP request for the identification of additional downriver hatchery sites, stating:

> "Additional candidate populations relevant to these SAR comparisons from downriver hatcheries of the Interior Columbia include Klickitat, Warm Springs, and Round Butte (depending on fish health constraints). Future monitoring should also consider incorporating PIT-tag SARs from the upper Columbia region to expand these regional comparisons."

In terms of additional wild downriver tag groups, Schaller et al. (2007; page 175) responded to the ISRP interest in complementing ongoing tagging in the John Day River by stating that the CSS

> "has proposed (but not received funding for) PIT-tagging wild spring Chinook smolts in the Warm Springs River (Deschutes Subbasin). Additional candidate populations relevant to these SAR comparisons from downriver areas of the Interior Columbia include Klickitat and Yakima rivers. Future monitoring should also consider incorporating PIT-tag SARs from the upper Columbia region to expand these regional comparisons."

Clearly, the CSS has identified several wild and hatchery population groups that could be used to quantify between-population variation in downriver areas. But the ISAB and ISRP have elected to terminate and not support the possibility of the CSS project collecting the data to monitor this variation.

These recommendations to terminate the Carson tagging effort and not support the proposed Warm Springs tagging effort are not compatible with key management information needs. The CSS monitors life-cycle survival rates and other demographic rates among upriver and downriver population groups within the Columbia Basin ecosystem. Good et al. (2007) describes the management importance of collecting these types of data at large spatial scales:

> "The strategy of the recovery planning process has thus been to confront the large-scale biological and management challenges by incorporating of relevant scientific information at similarly large scales."

Good et al. (2007) goes on to describe how the productivity and life-cycle survival data, such as those collected within the CSS, are important for the assessment of managed populations:

> "Stage-specific or lifetime productivity (i.e., population growth rate) provides information on important demographic processes. Abundance and productivity data are used to assess the status of populations of threatened and endangered ESUs."

Finally, when describing the management information needs of the Technical Recovery Teams (TRTs), Good et al. (2007) describes how large-scale observational data, including the natural variation in those data, is useful for quantifying cumulative effects:

> "The TRTs have moved on to analyses of the cumulative effects of multiple factors over large spatial scales, employing metapopulation models where possible, and fostering the use of large-scale experimentation to manipulate or take advantage of natural variation in ecological factors."

The information collected from downriver groups within the CSS can be used to help fulfill these critical management information needs. Therefore terminating the Carson tagging efforts and not supporting the Warm Springs tagging efforts is not compatible with key management information needs.

The scope of the CSS is primarily to establish a long-term dataset that measures the survival rate of annual generations of salmon and steelhead. Secondarily, the scope is to examine empirical evidence relating to explanatory hypotheses and to determine functional relationships useful for informing management decisions. In their Review, the ISAB and ISRP misinterpret the scope of the CSS as attempting to determine "unambiguous assignment of cause(s)" for all observed differences. These misinterpretations form the foundation for their rationale to terminate the existing, and failure to support the proposed, tagging efforts.

The primary objective of the CSS is to establish a long-term dataset that measures the survival rate of annual generations of salmon from their outmigration as smolts to their return to freshwater as adults to spawn (smolt-to-adult return rate; SAR). Secondarily, the scope of the CSS is to examine multiple hypotheses that attempt to explain the observed patterns of mortality within and between upriver and downriver stocks as well as hypotheses on the effects of freshwater and marine habitat conditions on the survival and migration rates of population groups. These secondary evaluations are conducted to determine functional relationships that may be useful for informing management decisions.

In their Review, the ISAB and ISRP misinterpret the scope of the CSS as attempting to determine "unambiguous assignment of cause(s)" for all observed differences between upriver and downriver population groups. None of the analyses in the CSS (or any other study in the region, including randomized experiments) are capable of proving cause-effect relationships (Rothman and Greenland 1998). To be clear, as stated in the CSS Report, "our specific interest through the CSS was whether upriver/downriver differences in overall survival for wild and/or hatchery stream-type Chinook (with more precise estimates from PIT-tagged groups) *were consistent with the differential mortality estimated from S-R models for wild populations*" (Schaller et al. 2007; page 7). These types of comparisons would also be useful for indentifying specific periods within the life-cycle when survival may be a limiting factor among population groups. When the ISAB and ISRP infer that the scope of the CSS is to provide "unambiguous assignment of cause(s)" for differences between population groups compared in the CSS, we feel that this is a misinterpretation of our scope. The ISAB and ISRP then go on to cite reasons why "unambiguous assignment of cause(s)" is impossible (reasons that we are well-aware of), and use these reasons as foundation and justification for eliminating the downriver tagging efforts. These "unambiguous assignment of cause(s)" criteria are unreasonable and unattainable for all of the observational studies in the region that are used to inform management decisions, including the CSS.

In addition to these issues associated with the elimination of monitoring tag groups, we would like to note several misconceptions apparent in the ISAB and ISRP interpretations of the CSS Report. The ISAB and ISRP comments are listed below, with our responses following.

Page 12: "The core reason a contrast of salmon survival between upriver and downriver locations is not advised is that the populations in tributaries downriver of the dams are not replicates of the upper Snake River populations."

Response: While the upriver/downriver populations compared within the CSS Report were not presented or characterized as perfect replicates of each other, making comparisons among population groups is a well-established and common scientific method for examining and drawing inferences about patterns in survival, productivity, and other demographic measures among salmonid population groups (see 39 publications listed above and recommendations of McDonald et al. (2007)). Nowhere do we state or imply that populations are perfect replicates of other populations. On the contrary, we consider each marked group as a replicate only of itself, and only in that year. The information obtained from each marked group may or may not share similar patterns with other marked groups. As stated above, the primary objective of the CSS is to measure the biological responses of population groups. Secondarily, we describe which marked groups are similar and which are different by making comparisons, and then go on to examine which factors are associated and which factors are unassociated with those similarities and differences.

Page 12: "Geographical variation in habitat types, productivity, predator populations, and local climatic conditions makes cause and effect interpretation problematic, even if more hatchery and downriver wild stocks could be identified."

Response: We clearly stated that the purpose (p. 106) of the upriver/downriver SAR comparisons was to determine if the previously-established differences in mortality estimated from spawner-recruit (SR) analyses were also apparent in the PIT-tag derived SARs. Nowhere did we state or imply that the purpose of the upriver/downriver comparisons was to determine cause and effect relationships; this is a false premise and a mischaracterization of the scope of the CSS Report. Contrasts of the point estimates and 90% CI from the two types of data (p. 131-133) indicated SAR-based estimates of differential mortality agreed well with published SR-based estimates of differential mortality. We characterized the upriver-downriver comparison as a "natural experiment", which therefore has more limited scope for inference than manipulative experiments (p. 150), but the approach and methods are consistent with observational studies in scientific literature (see 39 publications above). Further, we investigated and tested hypotheses regarding possible biological causes (including alternative hypotheses previously suggested by the NOAA NWFSC) of differential mortality between upriver and downriver wild stream-type Chinook (p. 136-143).

The CSS Report thoroughly investigated all of the testable scientific hypotheses that had been proposed as candidates for the SR-based estimates of differential mortality. The CSS Report performed carefully-constructed comparisons to examine the evidence for and against those hypotheses. Neither the content nor the outcomes of these comparisons were recognized in the Review. Instead, the ISAB and ISRP present newly proposed hypotheses and then criticize the CSS as being incapable of examining those hypotheses against existing data sets without allowing any opportunity for the CSS authors to investigate the evidence for or against those hypotheses.

Page 13: "The sponsors have presented evidence suggestive of a hydrosystem effect on differences in SARs between upriver and downriver sites, but little may be gained from further analysis of differences in SARs. The major conclusions of the research are already available for scrutiny by scientists and managers in peer-reviewed scientific literature and reports including the retrospective summary."

Response: Based on 5 years of PIT-tag SAR comparisons between wild Snake River and John Day River smolts, we have seen a consistent pattern of differential mortality across poor and favorable ocean conditions. Combined with estimates of in-river survival and relative survival of transported smolts, this is one line of (indirect) evidence that the magnitude of delayed hydrosystem mortality is large (e.g., Peters and Marmorek 2001; Schaller and Petrosky 2007). We strongly disagree with the opinion that little may be gained from further analyses of differences in SARs. On the contrary, we view the most critical task at hand as being to examine factors associated with those observed differences using life-stage-specific survival data, consistent with the recommendations of Good et al. (2007). We believe that these types of analyses, in conjunction with the detailed studies for fish travel time, reach survival, in-river SARs, transport SARs, and overall SARs are the critical measures that can and should be examined relative to candidate environmental and management factors. These passage characteristics and survival data provide critical information needed to inform management decisions regarding fish passage and survival through the FCRPS, and the likely effects of those decisions on adult returns (Peters and Marmorek 2001, Wilson 2003).

Page 9: "This is a single river system, without comparative measures of fish performance from before the hydrosystem was constructed, which makes unambiguous assignment of cause(s) impossible even if convincing, statistically significant differences in fish performance were established between upriver and downriver stocks."

Response: Contrary to this assertion, there are stock-recruitment data measuring comparative fish performances prior to and following full development of the FCRPS (e.g., John Day: 1959-1995; Imnaha: 1949-1995; Schaller et al. 1999). These stock-recruitment data have already established significant differences in fish performance between upriver (seven stocks) and downriver (six stocks) populations, and that the differences coincided with the construction of the hydrosystem (Schaller et al. 1999, Deriso et al. 2001). The task at hand is to examine factors currently associated with those observed stock-recruitment differences using life-stage-specific survival data (Good et al. 2007), and PIT-tag data provide the best opportunity to monitor life-stage-specific survival rates for making those comparisons.

Collection of the stock-recruitment data is an ongoing task among the fishery management agencies, and these data will provide ongoing opportunities to make comparisons with life-stage-specific survival data, but only if PIT-tag data are available. The primary limiting factor that is preventing further comparisons between stock-recruitment data with life-stage-specific survival data is the lack of downriver populations marked with PIT-tags. As mentioned above, nowhere does the CSS Report state or imply that the purpose of the comparisons is to assign causes, unambiguous or otherwise.

Page 13: "We now doubt that there are a sufficient number of appropriate downriver wild stocks available to make a meaningful comparison…All differences between upriver and downriver stocks would be candidates for causal factors, and, as we note above, it seems impossible to adequately control or rule out all alternative causes."

Response: In the Report, we identified Warm Springs, Klickitat, and Yakima Rivers as other wild downriver groups that would be useful for expanding downriver tagging efforts. We also identified the Warm Springs, Klickitat, and Round Butte hatchery stocks as useful additions for expanding downriver tagging efforts. It is an unreasonable and unachievable standard to require that all causative factors must be identified and accounted for prior to drawing any inferences on factors associated or unassociated with observed patterns of variation in observational studies. If the ISAB and ISRP believe that this standard must be met before making comparisons and drawing inferences, then all scientific investigations of real-world phenomena using observational data would fail to meet this criterion. This would include each of the 39 publications listed above. This would also include comparisons made by ISAB and ISRP members themselves using observational data (e.g., Poe et al. 1991; Bilby et al. 1998, 2003; Pearcy et al. 1996; Sork et al. 2005; Weeder et al. 2005). Clearly, the ISAB and ISRP are applying an arbitrary and unreasonable standard in judging the upriver/downriver comparisons presented in the CSS Report.

Page 9: "In sum, the system is too complex, and the possible sampling design necessarily too constrained in time and place, to reach conclusive findings on causation from this type of comparison."

Response: Making comparisons among salmon population groups across space and time, and subsequently drawing inferences based on the resulting observed patterns is a well-established and common scientific methodology (see 39 publications listed above). The level of complexity among stocks in the Columbia Basin is highly unlikely to be greater than the level of complexity among stocks separated by up to 3,000 km in the northeastern Pacific Ocean, as has been investigated by other researchers (Adkison et al. 1996; Botsford and Lawrence 2002; Bradford 1995; Brodeur et al. 2004, 2007; Fisher et al. 2007; Hare et al. 1999; Hodgson et al. 2006; Holt and Peterman 2004; Mantua et al. 1997; Morris et al. 2007; Mueter et al. 2002a, 2002b, 2005, 2007; Orsi et al. 2007; Peterman et al. 1998, 2003; Peterman 2004; Pyper et al. 1999, 2001, 2002, 2005; Pyper and Peterman 1999; Schindler et al. 2003; Su et al. 2004; Trudel et al. 2007a, 2007b; Quinn et al. 2005). Again, nowhere does the CSS Report state or imply that the scope of the comparisons is to conclusively find or determine causes of the observed differences.

Page 9: "This is a single river system, without comparative measures of fish performance from before the hydrosystem was constructed, which makes unambiguous assignment of cause(s) impossible even if convincing, statistically significant differences in fish performance were established between upriver and downriver stocks. In sum, the system is too complex, and the possible sampling design necessarily too constrained in time and place, to reach conclusive findings on causation from this type of comparison."
Pages 12-13: "There is inevitable confounding of all differences between downriver and upriver stocks and their environments, precluding clear attribution of cause for any upriver/downriver differences that might be shown."

Response: Significant differences between upstream and downstream SARs are found (e.g. non-overlapping confidence intervals in Figure 5.12). Unambiguous assignment of cause from this study alone is not necessary for the analysis to be useful. Other lines of evidence address the question of effects of the hydrosystem. There are comparative measures of SARs and estuary/ocean survival from before the completion of the hydrosystem (see Petrosky et al. 2001 and Wilson 2003).

Page 15: "Additional analyses related to differential mortality and survival metrics related to dam passage such as route of passage, temperature, cumulative stress, and predation are lacking."

Response: Chapter 4 compares SARs of C_0 and C_1, which addresses questions of route of passage and cumulative stress. Cumulative stress is a possible explanation for differences in SARs between upriver and downriver stocks seen in Chapter 5. Given the purpose, scope, and focus of the CSS, analyzing predation impacts is not practical; there are other studies which examine that issue.

Page 18: "The C_0 cohort remains undetected until it is below BON, and its numbers and survival through the hydrosystem must be estimated, inevitably more by assumption than by checkable results. That translates into a substantial estimation variance."

Response: Fish in the C_0 group are defined as undetected only at the collector projects (LGR, LGO, and LMN). A substantial portion is detected downriver at other projects

(MCN, JDA, BON, and the estuary PIT-trawl). The number of smolts in the C_0 category is estimated rather than observed, but estimation is done through standard CJS methodology and the accuracy of the estimation process is explored in Chapter 7 (and was also investigated indirectly in the 2002 Annual Report). The magnitude of estimation variance of number of C_0 smolts is detailed in Chapter 4, as well as Appendix E, and is usually quite small (1-6% CV). The number of returning adults in the C_0 category is not an estimate but a count of fish detected at LGR.

Page 19: "What is more bothersome is the fact that in the process of back-translating LGS- and LMN-transported stocks to LGR, we lose any opportunity to compare the SARs of these additional cohorts, call them T_1 and T_2 for convenience, with the SARs of T_0, C_0 and C_1. Such contrasts could be very useful indices of the success to be gained (or lost) by transportation from lower on the hydrosystem. ISAB asked for some resolution of this issue as early as the COMPASS (ISAB 2006-2) report, but this matter has yet to be clearly resolved."

Response: Transport project-specific SAR distributions are estimated and compared to in-river (C_0) SARs in Chapter 4. Given the limited annual sample sizes, and the fact that the LGR equivalent transport smolt number (T_0) is larger than the number of smolts transported from any project, reporting project-specific TIRs for individual years would be of limited value.

Page 19: "It seems unlikely that D will ever be definitively determined. Moreover, interpretation of D as the 'out of hydrosystem effect' requires that we separate the survival experience of BON-ocean-BON from that of BON-LGR, or at least the assumption that survival from BON-LGR is the same for T_0 as for C_0 (and/or for C_1). In spite of the small numbers returning for the upriver journey, it would be useful in the future to make an attempt to separate the effects."
Page 22: "We question, with reference to Chapter 3, whether it is reasonable to assume that D could be allocated to "out of hydrosystem effects," which at very least would require that the BON → spawning ground component of SAR(T_0) and SAR(C_0) could plausibly be as identical. This chapter clearly shows that assumption to be false; there are "within hydrosystem effects" embedded in D. There is unequivocal evidence that SAR(C_0) > SAR(T_0) for the BON-LGR leg of the upriver journey."

Response: There is no need for survival from BON-LGR of T_0 and C_0 or C_1 fish to be identical—surviving adults are enumerated at LGR. Any difference between groups is properly reflected in the ratio of SARs [TIRs or SAR(C_0) / SAR(C_1), or *D*]. *D* is intended to reflect effects *below the hydrosystem*, or "post-Bonneville", i.e. subsequent to the juvenile downstream migration up until returning to LGR as adults. Effects reflected in *D* (or TIR) do not necessarily have to be *outside the hydrosystem*, given that fish must migrate up the same hydrosystem they traversed downstream as juveniles. "Definitive determination" of *D*, like that of SARs, TIRs, or S_R is not possible. It is also not necessary in order to develop useful information for management decisions. These parameters are necessarily estimated with uncertainty; this uncertainty is reflected in the probability distributions presented and in the conclusions drawn.

Page 23: "Smolts transported from LGS (or from lower projects) fare better on the homeward journey than do those transported from LGR, but we have no reports on the compensating benefits on the outward journey, because the data on survival of T1, T2, and so on have been "back-adjusted to LGR" and are lumped with LGR-transported fish (T0). That is not necessary, and indeed, it is time for a direct assessment of SAR- and TIR-values for fish of the T1 (LGS-transported), T2 (LMN-transported), and other transported smolts as well, all in comparison with the C0 and T0 cohorts."

Response: Direct, project-specific estimates of transport SARs, TIRs, and *D*s were made for wild Chinook and wild steelhead in Chapter 4.

We hope that the ISAB/ISRP will receive this additional information in the spirit in which it is intended, of collaborative scientific exchange. The result of the ISAB/ISRP review comment was to limit tagging efforts and eliminate specific types of analysis. We believe that the censoring of analyses and prevention of data collection efforts, rather than constructive review and scientific criticism of analysis, does not advance our knowledge and tools to contribute to resolution of the current critical status of salmon and steelhead in the Columbia Basin.

Sincerely

Earl Weber CRITFC	Thomas Berggren FPC	Robin Ehlke WDFW
Eric Tinus ODFW	Tim Dalton ODFW	Charles Petrosky IDFG

Howard Schaller	Steve Haeseker	Paul Wilson
USFWS	USFWS	USFWS

References:

Adams, S.M. 2003. Establishing causality between environmental stressors and effects on aquatic ecosystems. Human and Ecological Risk Assessment 9(1):17-35.

Adkison, M.D., R.M. Peterman, M.F. Lapointe, D.M. Gillis, and J. Korman. 1996. Alternative models of climatic effects on sockeye salmon, Oncorhynchus nerka, productivity in Bristol Bay, Alaska, and the Fraser River, British Columbia. Fisheries Oceanography 5:137-152.

Beyers, D.W. 1998. Causal inference in environmental impact studies. Journal of the North American Benthological Society 17:367-373.

Bilby, R.E., B.R. Fransen, P.A. Bisson, and J.K. Walter. 1998. Response of juvenile coho salmon (Oncorhynchus kisutch) and steelhead (Oncorhynchus mykiss) to the addition of salmon carcasses to two streams in southwestern Washington, U.S.A. Canadian Journal of Fisheries and Aquatic Sciences 55:1909-1918.

Bilby R.E., E.W. Beach, B.R. Fransen, J.K. Walter, and P.A. Bisson. 2003. Transfer of nutrients from spawning salmon to riparian vegetation in western Washington. Transactions of the American Fisheries Society 132(4):733-745.

Botsford, L.W., and C.M. Paulsen. 2000. Assessing covariability among populations in the presence of intraseries correlation: Columbia River spring-summer chinook salmon (Oncorhynchus tshawytscha) stocks. Canadian Journal of Fisheries and Aquatic Sciences 57(3):616-627.

Botsford, L.W., and C.A. Lawrence. 2002. Patterns of co-variability among California Current Chinook salmon, coho salmon, Dungeness crab, and physical oceanographic conditions. Progress in Oceanography 53:283-305.

Bradford, M.J. 1995. Comparative review of Pacific salmon survival rates. Canadian Journal of Fisheries and Aquatic Sciences 52:1327-1338.

Brodeur, R.D., J.P. Fisher, D.J. Teel, R.L. Emmett, E. Casillas, and T.W. Miller. 2004. Juvenile salmonid distribution, growth, condition, origin, and environmental and species associations in the Northern California Current. Fishery Bulletin 102(1):25-46.

Brown, C.L., F. Parchaso, J.K. Thompson, and S.N. Luoma. 2003. Assessing toxicant effects in a complex estuary: a case study of effects of silver on reproduction in the bivalve, *Potamocorbula amurensis*, in San Francisco Bay. Human and Ecological Risk Assessment 9(1):95-119.

Budy, P., G.P. Thiede, N. Bouwes, C.E. Petrosky, and H. Schaller. 2002. Evidence linking delayed mortality of Snake River salmon to their earlier hydrosystem experience. North American Journal of Fisheries Management 22(1):35-51.

Budy, P., and H. Schaller. 2007. Evaluating tributary restoration potential for pacific salmon recovery. Ecological Applications 17(4):1068-1086.

Burnham, K.P., and D.R. Anderson. 2002. Model Selection and Multimodel Inference: a Practical Information-Theoretic Approach. Springer-Verlag, New York.

Bordeur, R.A., E.A. Daly, M.V. Sturdevant, T.W. Miller, J.H. Moss, M.E. Thiess, M. Trudel, L.A. Weitkamp, J. Armstrong, and E.C. Norton. 2007. Regional comparisons of juvenile salmon feeding in coastal marine waters off the west coast of North America. Pages 183-204 *in* C.B. Grimes, R.D. Brodeur, L.J. Haldorson, and S.M. McKinnell, editors. The ecology of juvenile salmon in the northeast Pacific Ocean: regional comparisons. American Fisheries Society, Symposium 57, Bethesda, Maryland.

Cochran, W.G. 1983. Planning and Analysis of Observational Studies. John Wiley & Sons, New York.

Collier, T.K. 2003. Forensic ecotoxicology: establishing causality between contaminants and biological effects in field studies. Human and Ecological Risk Assessment 9(1):259-266.

Chamberlin, T.C. 1890. The method of multiple working hypotheses. Science 15, 93 (Reprinted 1965, Science 148:754-759).

Culp, J.M., and D.J. Baird. 2006. Establishing cause-effect relationships in multi-stressor environments. *In* Methods in Stream Ecology. F.R. Hauer and G.A. Lamberti, editors. Elsevier, Oxford, United Kingdom.

Deriso, R., D. Marmorek, and I. Parnell. 1996. Retrospective analysis of passage mortality of spring Chinook of the Columbia River. *In* Plan for analyzing and testing hypotheses (PATH): final report of retrospective analysis for fiscal year 1996. *Compiled and edited by* D.R.Marmorek and 21 coauthors. ESSA Technologies Ltd., Vancouver, B.C.

Deriso, R.B., D.R. Marmorek, and I.J. Parnell. 2001. Retrospective patterns of differential mortality and common year-effects experienced by spring and summer chinook salmon (Oncorhynchus tshawytscha) of the Columbia River. Canadian Journal of Fisheries and Aquatic Sciences 58(12):2419-2430.

Diamond, J.M. 1986. Overview: laboratory experiments, field experiments, and natural experiments. *In* Community Ecology. J.M. Diamond and T.J. Case, editors. Harper & Row, New York.

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

Fisher, J., M. Trudel, A. Ammann, J.A. Orsi, J. Piccolo, C. Bucher, E. Casillas, J.A. Harding, R.B. Macfarlane, R.D. Brodeur, J.F.T. Morris, and D.W. Welch. 2007. Comparisons of the coastal distributions and abundances of juvenile Pacific salmon from central California to the northern Gulf of Alaska. Pages 31-80 *in* C.B. Grimes, R.D. Brodeur, L.J. Haldorson, and S.M. McKinnell, editors. The ecology of juvenile salmon in the northeast Pacific Ocean: regional comparisons. American Fisheries Society, Symposium 57, Bethesda, Maryland.

Fox, G.A. 1991. Practical causal inference for ecoepidemiologists. Journal of Toxicology and Environmental Health 33:359-373.

Gilbertson, M. 1997. Advances in forensic toxicology for establishing causality between Great Lakes epizootics and specific persistent toxic chemicals. Environmental Toxicology and Chemistry 16(9):1771-1778.

Good, T.P., T.J. Beechie, P. McElhany, M.M. McClure, and M.H. Ruckelshaus. 2007. Recovery planning for Endangered Species Act-listed Pacific salmon: using science to inform goals and strategies. Fisheries 32(9):426-440.

Hare, S.R., N.J. Mantua, and R.C. Francis. 1999. Inverse production regimes: Alaska and West Coast Pacific salmon. Fisheries 24(1):6-14.

Hilborn, R., and M. Mangel. 1997. The Ecological Detective: Confronting Models with Data. Princeton University Press, New Jersey.

Hodgson, S, T.P. Quinn, R. Hilborn, R.C. Francis, and D.E. Rogers. 2006. Marine and freshwater climatic factors affecting interannual variation in the timing of return migration to fresh water of sockeye salmon (Oncorhynchus nerka). Fisheries Oceanography 15(1):1-24.

Holt, C.A., and R.M. Peterman. 2004. Long-term trends in age-specific recruitment of sockeye salmon (Oncorhynchus nerka) in a changing environment. Canadian Journal of Fisheries and Aquatic Sciences 61(12):2455-2470.

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

ISAB. 2003-1. Review of Flow Augmentation: Update and Clarification.

ISAB. 2006-3. ISAB Review of the 2005 Comparative Survival Studies' Annual Report and Applicability of Comparative Survival Studies' Analysis Results.

ISRP & ISAB. 2005-15. Monitoring and Evaluation of Supplementation Projects.

ISAB & ISRP. 2007-6. Review of the Comparative Survival Study's (CSS) Ten-Year Retrospective Summary Report.

ISRP. 2004-8. Final Review of the United States Army Corps of Engineers' Anadromous Fish Evaluation Program for Fiscal Year 2004.

ISRP. 2005-14. Retrospective Report 1997 – 2005.

ISRP. 2006-6. Final Review of Proposals Submitted for Fiscal Years 2007-2009 Funding through the Columbia River Basin Fish and Wildlife Program.

ISRP. 2007-1. 2006 Retrospective Report.

ISRP. 2007-8. Review of John Day Study Plan for Project 2003-017-00, Integrated Status and Effectiveness Monitoring Program (ISEMP).

Jewel, N.P. 2005. Statistics for Epidemiology. Chapman & Hall/CRC, Boca Raton, Florida.

Lowell, R.B., J.M. Culp, and M.G. Dube. 2000. A weight-of-evidence approach for northern river risk assessment: integrating the effects of multiple stressors. Environmental Toxicology and Chemistry 19:1182-1190.

Magnusson, A, and R. Hilborn. 2003. Estuarine influence on survival rates of Coho (Oncorhynchus kisutch) and Chinook salmon (Oncorhynchus tshawytscha) released from hatcheries on the US Pacific Coast. Estuaries 26:1094-1103.

Mantua, N.J., S.R. Hare, Y. Zhang, J.M. Wallace, and R.C. Francis. 1997. A Pacific interdecadal climate oscillation with impacts on salmon production. Bulletin of the American Meteorological Society 78(6):1069-1079.

McDonald, L.L., R. Bilby, P.A. Bisson, C.C. Coutant, J.M. Epifanio, D. Goodman, S. Hanna, N. Huntley, E. Merrill, B. Riddell, W. Liss, E.J. Loudenslager, D.P. Philipp, W. Smoker, R.R. Whitney, and R.N. Williams. 2007. Research, monitoring, and evaluation of fish and wildlife restoration projects in the Columbia River Basin: lessons learned and suggestions for large-scale monitoring programs. Fisheries 32:582-590.

McHugh, P., P. Budy, and H. Schaller. 2004. A model-based assessment of the potential response of Snake River spring-summer Chinook salmon to habitat improvements. Transactions of the American Fisheries Society 133(3):622-638.

Morales, R., P. Gerhard, L. Andersson, J. Sturve, S. Rauch, and S. Molander. 2003. Establishing causality between exposure to metals and effects on fish. Human and Ecological Risk Assessment 9(1):149-169.

Morris, J.F.T., M. Trudel, M.E. Thiess, R.M. Sweeting, J. Fisher, S.A. Hinton, E.A. Fergusson, J.A. Orsi, E.V. Farley, Jr., and D.W. Welch. 2007. Stock-specific migrations of juvenile coho salmon derived from coded-wire tag recoveries on the Continental Shelf of western North American. Pages 81-104 *in* C.B. Grimes, R.D. Brodeur, L.J. Haldorson, and S.M. McKinnell, editors. The ecology of juvenile salmon in the northeast Pacific Ocean: regional comparisons. American Fisheries Society, Symposium 57, Bethesda, Maryland.

Mueter, F.J., R.M. Peterman, and B.J. Pyper. 2002a. Opposite effects of ocean temperature on survival rates of 120 stocks of Pacific salmon (Oncorhynchus spp.) in northern and southern areas. Canadian Journal of Fisheries and Aquatic Sciences 59(3):456-463.

Mueter, F.J., D.M. Ware, and R.M. Peterman. 2002b. Spatial correlation patterns in coastal environmental variables and survival rates of salmon in the north-east Pacific Ocean. Fisheries Oceanography 11(4):205-218.

Mueter, F.J., B.J. Pyper, and R.M. Peterman. 2005. Relationships between coastal ocean conditions and survival rates of northeast Pacific salmon at multiple lags. Transactions of the American Fisheries Society 134(1):105-119.

Mueter, F.J., J.L. Boldt, B.A. Megrey, and R.M. Peterman. 2007. Recruitment and survival of Northeast Pacific Ocean fish stocks: temporal trends, covariation, and regime shifts. Canadian Journal of Fisheries and Aquatic Sciences 64(6):911-927.

Orsi, J.A., J.A. Harding, S.S. Pool, R.D. Brodeur, L.J. Haldorson, J.M. Murphy, J.H. Moss, E.V. Farley, Jr., R.M. Sweeting, J.F.T. Morris, M.Trudel, R.J. Beamish, R.L. Emmett, and E.A. Fergusson. 2007. Epipelagic fish assemblages associated with juvenile Pacific salmon in neritic waters of the California Current and the Alaska Current. Pages 105-156 *in* C.B. Grimes, R.D. Brodeur, L.J. Haldorson, and S.M. McKinnell, editors. The ecology of juvenile salmon in the northeast Pacific Ocean: regional comparisons. American Fisheries Society, Symposium 57, Bethesda, Maryland.

Pearcy, W.G., J.P. Fisher, G. Anma, and T. Meguro. 1996. Species associations of epipelagic nekton of the North Pacific Ocean, 1978-1993. Fisheries Oceanography 5(1):1-20.

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 (Oncorhynchus nerka) stocks. Canadian Journal of Fisheries and Aquatic Sciences 55(11):2503-2517.

Peterman, R.M., B.J. Pyper, and B.W. MacGregor. 2003. Use of the Kalman filter to reconstruct historical trends in productivity of Bristol Bay sockeye salmon (Oncorhynchus nerka). Canadian Journal of Fisheries and Aquatic Sciences 60(7):809-824.

Peterman, R.M. 2004. Possible solutions to some challenges facing fisheries scientists and managers. ICES Journal of Marine Science 61(8):1331-1343.

Peters, C.N., and D.R. Marmorek. 2001. Application of decision analysis to evaluate recovery actions for threatened Snake River spring and summer Chinook salmon (Oncorhynchus tshawytscha). Canadian Journal of Fisheries and Aquatic Sciences 58(12):2431-2446.

Petrosky, C.E., and H.A. Schaller. 1992. A comparison of productivities for Snake River and Lower Columbia river spring and summer Chinook stocks. Proceedings of Salmon Managemetn in the 21st Century: Recovering Stocks in Decline. 1992 Northeast Pacific Chinook and Coho Workshop. American Fisheries Society, Idaho Chapter, Boise, Idaho.

Petrosky, C.E., and H.A. Schaller. 1996. Evaluation of productivity and survival rate trends in freshwater spawning and rearing life stage for Snake River spring and summer Chinook. *In* Plan for analyzing and testing hypotheses (PATH): final report of retrospective analysis for fiscal year 1996. *Compiled and edited by* D.R.Marmorek and 21 coauthors. ESSA Technologies Ltd., Vancouver, B.C.

Petrosky, C.E., H.A. Schaller, and P. Budy. 2001. Productivity and survival rate trends in the freshwater spawning and rearing stage of Snake River Chinook salmon (Oncorhynchus tshawytscha). Canadian Journal of Fisheries and Aquatic Sciences 58(6):1196-1207.

Poe, T.P., H.C. Hansel, S. Vigg, D.E. Palmer, and L.A. Prendergast. 1991. Feeding of predaceous fishes on out-migrating juvenile salmonids in John Day Reservoir, Columbia River. Transactions of the American Fisheries Society 120:405-420.

Pyper, B.J., and R.M. Peterman. 1999. Relationship among adult body length, abundance, and ocean temperature for British Columbia and Alaska sockeye salmon (Oncorhynchus nerka), 1967-1997. Canadian Journal of Fisheries and Aquatic Sciences 56(10):1716-1720.

Pyper, B.J., R.M. Peterman, M.F. Lapointe, and C.J. Walters. 1999. Patterns of covariation in length and age at maturity of British Columbia and Alaska sockeye salmon (Oncorhynchus nerka) stocks. Canadian Journal of Fisheries and Aquatic Sciences 56(6):1046-1057.

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(8):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.

Pyper, B.J., F.J. Mueter, and R.M. Peterman. 2005. Across-species comparisons of spatial scales of environmental effects on survival rates of Northeast Pacific salmon. Transactions of the American Fisheries Society 143(1):86-104.

Quinn, T.P., B.R. Dickerson, and L.A.Vollestad. 2005. Marine survival and distribution patterns of two Puget Sound hatchery populations of coho (Oncorhynchus kisutch) and chinook (Oncorhynchus tshawytscha) salmon. Fisheries Research 76(2):209-220.

Rasmussen, P.W., D.M. Heisey, E.V. Nordheim, and T.M. Frost. 2001. Time series intervention analysis: unreplicated large-scale experiments. *In* Design and Analysis of Ecological Experiments. S.M. Scheiner and J. Gurevitch, editors. Oxford University Press, Oxford, United Kingdom.

Roni, P., M.C. Liermann, C. Jordan, and E.A. Steel. 2005. Steps for designing a monitoring and evaluation program for aquatic restoration. Pages 13-34 *in* P. Roni, editor. Monitoring Stream and Watershed Restoration. American Fisheries Society, Bethesda, Maryland.

Rothman, K.J. and S. Greenland. 1998. Modern Epidemiology- Second Edition. Lippincott Williams & Wilkins, Philadelphia.

Schaller, H., C. Petrosky, and O. Langness. 1996. Contrasts in stock recruitment patterns of Snake and Columbia River spring and summer Chinook. *In* Plan for analyzing and testing hypotheses (PATH): final report of retrospective analysis for fiscal year 1996. *Compiled and edited by* D.R.Marmorek and 21 coauthors. ESSA Technologies Ltd., Vancouver, B.C.

Schaller, H.A., C.E. Petrosky, and O.P. Langness. 1999. Contrasting patterns of productivity and survival rates for stream-type Chinook salmon (Oncorhynchus tshawytscha) populations of the Snake and Columbia Rivers. Canadian Journal of Fisheries and Aquatic Sciences 56:1031-1045.

Schaller, H.A., and C.E. Petrosky. 2007. Assessing hydrosystem influence on delayed mortality of Snake River stream-type Chinook salmon. North American Journal of Fisheries Management 27(3):810-824.

Schaller, H., P. Wilson, S. Haeseker, C. Petrosky, E. Tinus, T. Dalton, R. Woodin, E. Weber, N. Bouwes, T. Berggren, J. McCann, S. Rassk, H. Franzoni, P. McHugh, M. DeHart. 2007. Comparative Survival Study (CSS) of PIT-Tagged Spring/Summer

Chinook and Steelhead in the Columbia River Basin: Ten-year Retrospective Summary Report. Project #1996-020-00 BPA Contract #s 25634, 25264, 20620.

Scheuerell, M.D. 2005. Influence of juvenile size on the age at maturity of individually marked wild Chinook salmon. Transactions of the American Fisheries Society 134(4):999-1004.

Schindler, D.E., M.D. Scheuerell, J.W. Moore, S.M. Gende, T.B. Francis, and W.J. Palen. 2003. Pacific salmon and the ecology of coastal ecosystems. Frontiers in Ecology and the Environment 1(1):31-37.

Seeb, L.W., A. Antonovich, M.A. Banks, T.D. Beacham, M.R. Bellinger, S.M. Blankenship, M.R. Campbell, N.A. Decovich, J.C. Garza, C.M. Guthrie III, T.A. Lundrigan, P. Moran, S.R. Narum, J.J. Stephenson, K.J. Supernault, D.J. Teel, W.D. Templin, J.K. Wenburg, S.F. Young, and C.T. Smith. 2007. Development of a standardized DNA database for Chinook salmon. Fisheries 32:540-549.

Su, Z.M., R.M. Peterman, and S.L. Haeseker. 2004. Spatial hierarchical Bayesian models for stock-recruitment analysis of pink salmon (Oncorhynchus gorbuscha) Canadian Journal of Fisheries and Aquatic Sciences 61(12):2471-2486.

Sork, V.L., P.E. Smouse, V.J. Apsit, R.J. Dyer, and R.D. Westfall. 2005. A two-generation analysis of pollen pool genetic structure in flowering dogwood, Cornus florida (Cornaceae), in the Missouri Ozarks. American Journal of Botany 92:262-271.

Trudel, M., S.R.M. Jones, M.E. Thiess, J.F.T. Morris, D.W. Welch, R.M. Sweeting, J.H. Moss, B.L. Wing, E.V. Farley, Jr., J.M. Murphy, R.E. Baldwin, and K.C. Jacobson. 2007a. Infestations of motile salmon lice on Pacific salmon along the west coast of North America. Pages 157-182 *in* C.B. Grimes, R.D. Brodeur, L.J. Haldorson, and S.M. McKinnell, editors. The ecology of juvenile salmon in the northeast Pacific Ocean: regional comparisons. American Fisheries Society, Symposium 57, Bethesda, Maryland.

Trudel, M., M.E. Thiess, C. Bucher, E.V. Farley, Jr., R.B. MacFarlane, E. Casillas, J. Fisher, J.F.T. Morris, J.M. Murphy, and D.W. Welch. 2007b. Regional variation in the marine growth and energy accumulation of juvenile Chinook salmon and coho salmon along the west coast of North America. Pages 205-232 *in* C.B. Grimes, R.D. Brodeur, L.J. Haldorson, and S.M. McKinnell, editors. The ecology of juvenile salmon in the northeast Pacific Ocean: regional comparisons. American Fisheries Society, Symposium 57, Bethesda, Maryland.

Weeder J.A., A.R. Marshall, and J.M. Epifanio. 2005. An assessment of population genetic variation in Chinook salmon from seven Michigan rivers 30 years after introduction. North American Journal of Fisheries Management 25:861-865.

Wilson, P.H. 2003. Using population projection matrices to evaluate recovery strategies for Snake River spring and summer Chinook salmon. Conservation Biology 17(3):782-794.

Woodward, M. 2005. Epidemiology: Study Design and Data Analysis. Chapman & Hall/CRC, Boca Raton, Florida.

Zabel, R.W., and S. Achord. 2004. Relating size of juveniles to survival within and among populations of Chinook salmon. Ecology 85(3):795-806.