Independent Scientific Advisory Board
for the Northwest Power & Conservation Council, Columbia River Basin Indian Tribes, and NOAA Fisheries

Independent Scientific Review Panel
for the Northwest Power & Conservation Council; 851 SW 6th Avenue, Suite 1100; Portland, Oregon 97204

Review of the Comparative Survival Study's (CSS) Ten-Year Retrospective Summary Report

ISAB & ISRP 2007-6
November 19, 2007

ISAB and ISRP Reviewers

Richard Alldredge, Ph.D., Professor of Statistics at Washington State University.

Robert Bilby, Ph.D., Ecologist at Weyerhaeuser Company.

Peter A. Bisson, Ph.D., Senior Scientist at the Olympia (Washington) Forestry Sciences Laboratory of the U.S. Forest Service.

John Epifanio, Ph.D., Director and Associate Professional Scientist for the Center for Aquatic Ecology at the Illinois Natural History Survey.

Linda Hardesty, Ph.D., Associate Professor of Range Management at Washington State University.

Charles Henny, Ph.D., Senior Research Scientist at the U.S. Geological Survey in Corvallis, Oregon.

Nancy Huntly, Ph.D., Professor of Wildlife Biology at Idaho State University.

Stuart Hurlbert, Ph.D., Professor of Biology and Director, Center for Inland Waters at San Diego State University.

Roland Lamberson, Ph.D., Professor of Mathematics and Director of Environmental Systems Graduate Program at Humboldt State University.

Colin Levings, Ph.D., Scientist Emeritus and Sessional Researcher, Department of Fisheries and Oceans. Canada

Eric J. Loudenslager, Ph.D., Hatchery Manager at Humboldt State University, California.

Katherine Myers, Ph.D., Principal Investigator of the High Seas Salmon Research Program at the School of Aquatic and Fishery Sciences, University of Washington.

William Pearcy, Ph.D., Professor Emeritus of Oceanography at Oregon State University.

Thomas P. Poe, M.S., Consulting Fisheries Scientist, formerly with the U.S. Geological Survey.

Dennis Scarnecchia, Ph. D., Professor of Fish and Wildlife Resources, University of Idaho.

Peter Smouse, Ph.D., Professor of Ecology, Evolution, and Natural Resources at Rutgers University.

Bruce Ward, Fisheries Scientist, Ministry Of Environment, Aquatic Ecosystem Science Section, U.B.C., Vancouver.

Chris Wood, Ph.D., Head, Conservation Biology Section, Department of Fisheries and Oceans, Canada.

Review of the Comparative Survival Study's Ten-year Retrospective Summary Report

Contents

Review of the Comparative Survival Study's Ten-year Retrospective Summary Report

Executive Summary

This report is the most recent in a series of ISAB and ISRP reviews of the Comparative Survival Study. The Northwest Power and Conservation Council (the Council) requested this current ISAB and ISRP review of the CSS Ten-Year Retrospective Summary Report[1] to inform the funding decision for the CSS proposal for Fiscal Years 2008 and 2009. This review follows on an earlier review by the ISAB of the CSS 2005 Annual Report, also requested by the Council, in which two questions were posed. These two questions were given provisional answers in the ISAB review, and a number of specific concerns were identified that made the Annual Report an inadequate source of information to answer those questions more thoroughly. The ISAB review of the CSS 2005 Annual Report (ISAB 2006-3[2]) included a recommendation that the CSS team prepare a summary and retrospective synthesis of the first 10 years of the project, because such a synthetic review of information and interpretation was needed to provide clear answers to the questions posed by the Council, as well as to support other management applications and scientific interpretations of the CSS results. This current review is of the resultant CSS 10-Year Retrospective Report, which has been completed in response to the ISAB 2006-3 recommendations and directive from the Council. This ISAB and ISRP review was requested by the Council, which also asked that the ISAB and ISRP evaluate the responsiveness of the Retrospective to comments in ISAB 2006-3 and again provide answers to the Council's two questions.

The ISAB and ISRP find that the ten-year summary report is clear, thorough, responsive to past ISAB comments, and was completed in a retrospective style, a major accomplishment for which we commend the CSS investigators. The ISAB/ISRP provide their detailed evaluation in four parts: the ISRP recommendation for the CSS FY 2007-09 proposal, the ISAB/ISRP response to the two questions posed by the Council in their 2005 request for review of the CSS, an evaluation of the effectiveness of the CSS Retrospective in answering the concerns that were posed by the ISAB's review of the 2005 Annual Report, and chapter by chapter specific comments for the CSS team.

The ISRP finds that the CSS FY 2007-09 proposal Meets Scientific Review Criteria (In Part). Specifically, the ISRP finds that the first three biological objectives of the CSS proposal (Estimate Smolt to Adult Survival Rates [SARs], SAR Hydro Goal, and Transport to Control [T/C] Ratios) meet scientific review criteria. The ISRP finds that the fourth objective (Upriver/Downriver Comparisons) does not meet scientific review criteria, because of inevitable confounding from other factors in establishing cause(s) of upriver/downriver differences that may be detected, regardless of sample size and detection power that could be achieved.

---

[1] www.fpc.org/documents/CSS.html

[2] ISAB Review of the 2005 Comparative Survival Studies' Annual Report and Applicability of Comparative Survival Studies' Analysis Results: www.nwcouncil.org/library/isab/isab2006-3.htm

Overall, the CSS Ten-Year Retrospective was effective in answering the concerns posed by the ISAB's review of the CSS 2005 Annual Report (ISAB 2006-3). The Retrospective provided improved clarity in the presentation and explanation of the sophisticated methodologies used in analyses of CSS data. The scope of CSS investigations resulted in an extensive report, containing many detailed summaries of past and present work, and the report presents key data and data summaries in support of their major conclusions. The CSS team has responded very well in a short time frame to the difficult challenge of including enough details to allow scientific review, while avoiding obfuscation by sheer volume of material.

The 10-Year Retrospective facilitated improved answers to the questions posed earlier by the Council:

> *Council Question 1. Are the design, implementation, and interpretation of the statistical analyses underpinning the report based on the best available methods? Does the ISAB have suggestions for improving the analyses?*

Similar to the ISAB review of the 2005 CSS Annual Report, this current ISAB/ISRP review finds that the design, implementation, and interpretation underpinning the 10-Year Retrospective Report are very good. The CSS constitutes a successful implementation of a large-scale tagging program. The CSS has benefited from also using PIT-tags from other marking programs, and we encourage even more cooperation among PIT-tag marking programs to address critical uncertainties and improve reliability of survival estimates. We have included advice in the detailed chapter reviews for the CSS staff on design and analysis issues. We have also noted where we believe other interpretations of results should be considered or altered, e.g., the upriver/downriver comparative analyses.

> *Council Question 2. What is the applicability of the CSS results, taking into account whatever scientific criticisms of the analyses that the ISAB decides are valid, if any? In other words, what weight should the analyses be given and what qualifiers should be considered when using the analyses for decision-making?*

The ISAB and ISRP find that the CSS results are based upon carefully considered and applied methods of analysis. It is inevitable that there are other methods of analysis and that there can be other interpretations of results. We support the CSS efforts to refine analytical methodology, analyze other data, and design additional studies to collect more data to answer important questions for the region.

Caution is always needed in interpreting results, and the assumptions that are used in interpretation, as well as measures of uncertainty, must be taken into account in deciding the application of any interpretation. For instance, current conclusions that transportation provided, or did not provide, benefit to a species or wild/hatchery group requires qualification with the possibility of selection bias of fish for transportation due to size, condition, location in the water column, etc. Similarly, conclusions about mortality or delayed mortality of transported fish, relative to in-river fish, are not equivalent to saying that mortality or delayed mortality are DUE to transportation, <u>unless</u> all other factors can be discounted. Similarly, statements that trends are consistent with a specific hypothesis are most useful where alternative explanations are examined and discounted.

Both the ISRP and ISAB agree that the upriver/downriver comparative analyses, with ensuing inferences of causation, should be discontinued. Although the basic data on performance of upriver and downriver stocks remain of value in monitoring and evaluation, inevitable confounding in the sampling design precludes unambiguous interpretation of cause of upriver/downriver differences. That is not to say that upriver/downriver differences are absent, but rather methods to identify causality are lacking.

We also find many well-supported interpretations in the CSS Retrospective that should be carefully considered by Council and other decision-makers. These are discussed in detail in later sections of this report, particularly V.B, V.D, V.E, and V.G. We note several here, referencing the sections of the report where they are presented in full detail.

The CSS report suggests that the net benefit of transportation decreases in the order (hatchery and wild Steelhead) > (hatchery Chinook) > (wild Chinook), with wild Chinook showing no consistent evidence of any net benefit. Sample sizes of hatchery Chinook are large enough that the evidence for $TIR_0 > 1$ (transport to in-river ratio) was collectively convincing. Transportation of steelhead seems to be beneficial enough that even limited precision is not a serious impediment to establishing that fact. For wild Chinook, the sample size is inadequate and nothing very convincing can be said. Even where there was substantial benefit from transportation, none of these SARs were sufficient for stock persistence, so while it sometimes did improve survival, transportation alone was not sufficient to ensure persistence. (See Chapter 3 of the CSS Retrospective, devoted to a comparison of in-river and transported fish, where the sampling unit of interest is the entire yearly run, and section V.B of this ISAB/ISRP review.)

The report finds that overall SARs for wild spring/summer Chinook (geometric mean 0.9%, range 0.3-2.4%) and wild steelhead (geometric mean 1.6%, range 0.3% - 2.9%) fell short of the NPCC SAR recovery goals (2% minimum, 4% average). These conclusions were supported by the data presented in the report but do not include the untagged fish data; for some years, the untagged fish have been estimated to survive at considerably higher rates than tagged fish (Copeland et al. 2007). Regardless, these SARs are low and are reason for concern for recovery of these listed stocks. (See Chapter 5 of the Retrospective, evaluating and comparing SARs for Snake River listed salmon and steelhead, and section V.D. of this review.)

The report indicates that transportation, whatever benefits it may have on the outward journey, has negative consequences on the homeward journey, and they are substantial. Smolts transported from Lower Goose Dam (or from lower projects) fare better on the homeward journey than do those transported from Lower Granite Dam. However, we have no reports on the compensating benefits on the outward journey. We need direct assessment of SAR- and TIR-values for fish of the T1 and other transported smolts, all in comparison with the C0 and T0 cohorts. We need proper TIR-values for all of the sampled cohorts, measured Lower Granite Dam to Lower Granite Dam. The retrospective report reminds us that the Lower Granite Dam to natal source component of survival is outside the current CSS mandate, but suggests that we need to know more about it. We concur. (See Chapter 6 of the CSS Retrospective, designed to evaluate survival of hatchery and wild Chinook on their return journey from Bonneville to Lower Granite Dam (LGR) and from LGR to their source hatchery or spawning ground, and section V.E of this ISAB/ISRP review.)

Recommendations

From review of the CSS Retrospective, the ISAB and ISRP suggest three general priorities for future work:

1. Initiate a comprehensive study to determine why the PIT tagged Snake River wild spring/summer Chinook are producing lower SARs than the unmarked wild Chinook.

2. Initiate a study to determine why wild spring-summer Chinook gain no benefit from transportation (TIR~1.0) compared to hatchery Chinook and steelhead.

3. Prepare and submit for peer-reviewed publication a major synthesis paper, highlighting central results and interpretations of the CSS study.

I. Introduction

This joint Independent Scientific Advisory Board (ISAB) and Independent Scientific Review Panel (ISRP) review of the Comparative Survival Study's (CSS) Ten-year Retrospective Summary Report reflects over a decade of iterative independent scientific review by the ISAB and ISRP, as well as adaptive study design and implementation, and subsequent reporting by the CSS project sponsors. The ISAB and ISRP conducted this review jointly because the CSS Ten-year Retrospective Summary Report (CSS Retrospective) presents findings with program-level applicability on hydrosystem operations suited for an ISAB review, while at the same time the CSS project is funded through the Council's Fish and Wildlife Program and is thus reviewed by the ISRP.

The CSS is a field study of the survival of PIT-tagged spring/summer Chinook and PIT-tagged summer steelhead through the Snake and Columbia River(s) hydrosystem from smolts through returning adults, with a focus on relative survival of fish that traveled as smolts by alternative routes (e.g., in river, transported, different routes of dam passage, and different numbers of dams passed).

The CSS is important because it is one of the few organized attempts to systematically release PIT-tagged hatchery-reared and wild smolts into the Columbia River for the purpose of comparative monitoring and evaluation. Most aspects of the study, from its design and methods to the analytical results, continue to be strongly debated in the Region because the relative survival rates of salmonids under different hydrosystem operations and environmental constraints is a central concern of water and fish management policies.

The Council requested this ISAB and ISRP review to inform the Council's funding decision on the CSS project for Fiscal Years (FY) 2008 and 2009. Specifically, as described in a November 21, 2006, letter from the Council to the Bonneville Power Administration, the Council recommended FY 2007 funding to "continue the ongoing level of PIT-tagging of spring/summer Chinook salmon and steelhead, in order to continue to obtain the PIT-tag data from the marked fish, which various projects rely upon and utilize." The Council recommended no funding to tag new groups. The Council also recommended funding to complete a retrospective summary report, as recommended by the ISAB and ISRP, and have it reviewed within the region and by the ISAB. With this ISAB/ISRP review in hand, the Council intends to make a funding recommendation to Bonneville for FY 2008 and 2009.

As per the Council's request, the CSS Retrospective was created in response to recommendations contained in the ISAB's review of the CSS's 2005 Annual Report. These ISAB recommendations were subsequently incorporated in the ISRP's review of the FY 2007-09 CSS proposal.[3] These were the latest in a series of iterative reviews of the CSS project over the past decade; see Appendix A (www.nwcouncil.org/library/isab/isabisrp2007-6app.pdf) for excerpts from past ISAB and ISRP reviews. This ISAB and ISRP review considers how well the CSS

[3] ISRP Final Review of Proposals submitted for Fiscal Years 2007-2009 Funding through the Columbia River Basin Fish and Wildlife Program: www.nwcouncil.org/library/isrp/isrp2006-6.htm

Retrospective addresses the ISAB and ISRP's recommendations from the review of the CSS's 2005 Annual Report and the Council's original questions posed to the ISAB. An updated ISRP recommendation for the CSS FY 2007-09 proposal (199602000) is also provided.

On December 20, 2005, the Council requested that the Independent Scientific Advisory Board (ISAB) review the 2005 Annual Report for the Comparative Survival Study (CSS) prepared by the Fish Passage Center (FPC) and the Comparative Survival Study Oversight Committee, as well as critical comments on the draft of that report by the Bonneville Power Administration (BPA) and NOAA Fisheries. The Council asked that the ISAB address the following specific questions:

> 1. Are the design, implementation, and interpretation of the statistical analyses underpinning the report based on the best available methods? Does the ISAB have suggestions for improving the analyses?
>
> 2. What is the applicability of the CSS results, taking into account whatever scientific criticisms of the analyses that the ISAB decides are valid, if any? In other words, what weight should the analyses be given and what qualifiers should be considered when using the analyses for decision-making?

With regard to Council question 1, the ISAB review made recommendations for improvement, while stating that the design, implementation, and interpretation underpinning the 2005 Annual Report were very good. The ISAB review noted that Council question 2 was difficult to answer with the present annual progress report and recommended that the CSS team produce a ten-year summary report including some additional in-depth interpretations and analyses of the data, in a retrospective style.

The ten-year summary report has been completed in a retrospective form. Producing this report was a major accomplishment, and we commend the CSS investigators for their efforts. The current ISAB/ISRP review is presented in four parts:

- ISRP Recommendation for the CSS FY 2007-09 proposal
- Response to Council questions
- Evaluation of the effectiveness of the CSS Retrospective in answering concerns raised by the ISAB review of the CSS 2005 Annual Report (ISAB 2006-3).
- Chapter by chapter specific comments for the CSS team

To complete this review, the ISAB and ISRP requested and received a very informative briefing on CSS Retrospective by the project sponsors at the ISAB and ISRP's September meeting. This briefing was followed by an informal question and answer with the project sponsor at the ISAB and ISRP's October meeting. The ISAB and ISRP reviewed the CSS Retrospective Report in depth, but only briefly reviewed the comments from others on the CSS report in the context of our chapter reviews. We do not address the comments point by point, and we did not re-analyze any of the CSS or commenters' specific data analyses due to the short time available for our review.

II. ISRP Recommendation for the CSS FY 2007-09 proposal

The ISRP concludes that CSS proposal #199602000: Meets Scientific Review Criteria (In Part)

Proposal Background

The FY 2007-09 CSS proposal identified four biological objectives.

1. Estimate SARs. Estimate smolt-to-adult survival rate (SAR) for transported wild and hatchery stream type Chinook and steelhead.
2. SAR Hydro Goal. Determine if SAR rates are significantly different from the interim SAR hydro goal.
3. T/C Ratios. Estimate transport/control ratio and in-river survival rates for wild and hatchery yearling Chinook and steelhead concurrently over a number of years in order to span a range of environmental conditions.
4. Upriver/Downriver Comparisons. Compare SARs of transported and downriver indicator stocks.

In the FY 2007-09 proposal these tasks would be accomplished by PIT tagging steelhead and Chinook in Tables 1 and 2.

Table 1. Number of hatchery steelhead (see tributary allocations on next page), hatchery Chinook, and additional wild Chinook smolts to be PIT tagged specifically for CSS.

Organi-zation	Budget Contacts	Tagging Site	Species and rearing type	PIT tag quota
IDFG	S. Kiefer	Magic Valley, Hagerman,	H Steelhead	(50,000 LSRCP)
	R. Duke	Clearwater, and Niagara Springs	H Steelhead	30,000 BPA
	E. Buettner	Rapid R Hatchery	H Chinook	52,000
		McCall Hatchery	H Chinook	52,000
		Salmon R Trap	W Chinook	5,000[A]
		Snake R Trap	W Chinook	2,000[A]
		Clearwater R Trap	W Chinook	3,200
		Clearwater R Trap	W Steelhead	1,400
		Other IDFG tributary traps	W Chinook	14,500[B]
ODFW	R. Carmichael	Irrigon Hatchery	H Steelhead	20,000
		Lookingglass Hatchery		
		• Imnaha R AP release	H Chinook	21,000[C]
		• Catherine Ck AP release	H Chinook	21,000[C]
		Grande Ronde R trap	W Chinook	1,400[A]
USFWS	D. Wills	Dworshak Hatchery	H Steelhead	25,000
	H. Burge	Dworshak Hatchery	H Chinook	52,000
		Carson Hatchery	H Chinook	15,000

[A] Additional smolts to be PIT tagged above the current SMP tagging quota levels.
[B] Cost for PIT tags only to complement on-going PIT tagging efforts in Idaho.
[C] Fish PIT tagged in the fall of the contract year for the next year's migration.

In addition, ODFW, under the project *Salmonid Productivity, Escapement, Trend, and Habitat Monitoring in the John Day River Subbasin* (199801600), is anticipated to PIT tag 6,000 wild Chinook and 6,000 wild steelhead in the John Day River.

In a letter from Michele DeHart (Fish Passage Center) to Tracy Hauser (BPA) September 24, 2007 it was indicated the CSS wished to PIT tag the following groups:

Table 2. Number of wild and hatchery Chinook and steelhead to be PIT tagged under CSS contract in 2008.

Organization	Tagging Site	Species and rearing type	PIT tags needed	Date tags Needed
IDFG	Magic Valley Hatchery Hagerman NFH Clearwater Hatchery Niagara Springs Hatchery Rapid R Hatchery McCall Hatchery Salmon R. Trap Snake R. Trap Clearwater R. Trap Clearwater R. Trap Other IDFG tributary traps	H Steelhead H Steelhead H Steelhead H Steelhead H Chinook H Chinook W Chinook W Chinook W Chinook W Steelhead W Chinook	13,134[A] 7,666[A] 5,200[A] 28,000 52,000 52,000 5,000 2,000 3,200 2,000 14,500[B]	1/1/08 1/1/08 1/1/08 1/1/08 1/15/08 1/15/08 3/5/08 3/5/08 3/5/08 3/5/08 2/6/08
ODFW	Irrigon Hatchery Lookingglass Hatchery • Imnaha R. release • Catherine Ck. release Grande Ronde R. Trap John Day River	H Steelhead H Chinook H Chinook W Chinook W Chinook W Steelhead	13,000[A] 21,000[C] 21,000[C] 1,400 6,000[D] 6,000[D]	12/1/07 9/3/08 9/3/08 3/5/08
USFWS	Dworshak NFH Dworshak NFH Carson NFH	H Steelhead H Chinook H Chinook	8,000[A] 52,000 15,000	12/12/07 12/12/07 12/12/07
Warms Spring Tribe	Warms Springs R. Trap (Deschutes River basin)	W Chinook	6,000	2/20/08

[A] Fish PIT tagged under CSS contract to complement the LSRCP's proposed steelhead hatchery evaluation (tagging) study.

[B] Cost for PIT tags only to complement on-going wild Chinook PIT tagging efforts in Idaho.

[C] Fish to be PIT tagged in September 2008 for the 2009 migration.

[D] Fish PIT tagged under CSS contract only if not renewed under existing ODFW contract with BPA.

This includes additional marking of wild steelhead, hatchery A and B run steelhead from the Snake River Basin, Warm Springs River (Deschutes subbasin) wild spring Chinook, and John Day wild steelhead. If ODFW's John Day monitoring and evaluation project (199801600) did not receive funding for PIT tagging 6,000 each wild spring Chinook and steelhead from the John Day, these would be added to the project request as well.

ISRP Comments

Biological objectives (1), (2), and (3) and associated work elements meet scientific review criteria – estimate SARs, SAR hydro goals, and T/C ratios. The upriver/downriver comparison, biological objective (4), and associated tagging and tasks do not meet scientific review criteria. 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. 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.

The ISRP acknowledges this is a departure from earlier ISRP reviews that identified the comparison as flawed with only Carson hatchery and John Day representing lower river stocks, but recommended expanded lower river tagging sites (ISRP 2006-6). The ISRP has reached this conclusion and altered recommendation after significant internal discussion based on the evaluation of the CSS Retrospective Report and from findings reported in the ISAB's Latent Mortality Report (ISAB 2007-1). The ISAB and ISRP's full rationale for this recommendation is provided below in sections IV. and V.D. under comments on Chapter 5 of the CSS Retrospective.

Because the upriver/downriver comparison is determined not to meet scientific review criteria, existing tagging at Carson National Fish Hatchery (15000 yearling Chinook salmon) and proposed expanded wild spring Chinook tagging in the Warm Springs River are not justified under proposal 199602000. Existing tagging of 6,000 wild spring Chinook and 6,000 wild steelhead in the John Day River by ODFW's John Day monitoring and evaluation project (199801600) is not justified as a contribution to the CSS project (199602000); however, this tagging effort may be justified for other evaluation purposes.

Expanded tagging of hatchery steelhead in the Snake River subbasin in cooperation with the Lower Snake River Compensation Plan appears justified to improve the estimates of metrics to accomplish biological objectives 1 – 3.

III. Response to Council Questions

As described above, for the ISAB's review of the CSS 2005 Annual Report, the Council asked that the ISAB address two specific questions:

> *1. Are the design, implementation, and interpretation of the statistical analyses underpinning the report based on the best available methods? Does the ISAB have suggestions for improving the analyses?*

Similar to the ISAB review of the 2005 CSS Annual Report, this current ISAB/ISRP review finds that the design, implementation, and interpretation underpinning the 10-Year Retrospective Report are very good. The CSS constitutes a successful implementation of a large-scale tagging program. The CSS has benefited from also using PIT-tags from other marking programs, and we encourage even more cooperation among PIT-tag marking programs to address critical uncertainties and improve reliability of survival estimates. We have included advice in the detailed chapter reviews for the CSS staff on design and analysis issues. We have also noted where we believe other interpretations of results should be considered or altered, e.g., the upriver/downriver comparative analyses.

> *2. What is the applicability of the CSS results, taking into account whatever scientific criticisms of the analyses that the ISAB decides are valid, if any? In other words, what weight should the analyses be given and what qualifiers should be considered when using the analyses for decision-making?*

The ISAB and ISRP find that the CSS results are based upon carefully considered and applied methods of analysis. It is inevitable that there are other methods of analysis and that there can be other interpretations of results. We support the CSS efforts to refine analytical methodology, analyze other data, and design additional studies to collect more data to answer important questions for the region.

Caution is always needed in interpreting results, and the assumptions that are used in interpretation, as well as measures of uncertainty, must be taken into account in deciding the application of any interpretation. For instance, current conclusions that transportation provided, or did not provide, benefit to a species or wild/hatchery group requires qualification with the possibility of selection bias of fish for transportation due to size, condition, location in the water column, etc. Similarly, conclusions about mortality or delayed mortality of transported fish, relative to in-river fish, are not equivalent to saying that mortality or delayed mortality are DUE to transportation, unless all other factors can be discounted. Similarly, statements that trends are consistent with a specific hypothesis are most useful where alternative explanations are examined and discounted.

Both the ISRP and ISAB agree that the upriver/downriver comparative analyses, with ensuing inferences of causation, should be discontinued. Although the basic data on performance of upriver and downriver stocks remain of value in monitoring and evaluation, inevitable confounding in the sampling design precludes unambiguous interpretation of cause of upriver/downriver differences. That is not to say that upriver/downriver differences are absent, but rather methods to identify causality are lacking.

We also find many well-supported interpretations in the CSS Retrospective that should be carefully considered by Council and other decision-makers. These are discussed in detail in later sections of this report, particularly V.B, V.D, V.E, and V.G. We note several here, referencing the sections of the report where they are presented in full detail.

The CSS report suggests that the net benefit of transportation decreases in the order (hatchery and wild Steelhead) > (hatchery Chinook) > (wild Chinook), with wild Chinook showing no consistent evidence of any net benefit. Sample sizes of hatchery Chinook are large enough that the evidence for $TIR_0 > 1$ (transport to in-river ratio) was collectively convincing. Transportation of steelhead seems to be beneficial enough that even limited precision is not a serious impediment to establishing that fact. For wild Chinook, the sample size is inadequate and nothing very convincing can be said. Even where there was substantial benefit from transportation, none of these SARs were sufficient for stock persistence, so while it sometimes did improve survival, transportation alone was not sufficient to ensure persistence. (See Chapter 3 of the CSS Retrospective, devoted to a comparison of in-river and transported fish, where the sampling unit of interest is the entire yearly run, and section V.B of this ISAB/ISRP review.)

The report finds that overall SARs for wild spring/summer Chinook (geometric mean 0.9%, range 0.3-2.4%) and wild steelhead (geometric mean 1.6%, range 0.3% - 2.9%) fell short of the NPCC SAR recovery goals (2% minimum, 4% average). These conclusions were supported by the data presented in the report but do not include the untagged fish data; for some years, the untagged fish have been estimated to survive at considerably higher rates than tagged fish (Copeland et al. 2007). Regardless, these SARs are low and are reason for concern for recovery of these listed stocks. (See Chapter 5 of the Retrospective, evaluating and comparing SARs for Snake River listed salmon and steelhead, and section V.D. of this review.)

The report indicates that transportation, whatever benefits it may have on the outward journey, has negative consequences on the homeward journey, and they are substantial. Smolts transported from Lower Goose Dam (or from lower projects) fare better on the homeward journey than do those transported from Lower Granite Dam. However, we have no reports on the compensating benefits on the outward journey. We need direct assessment of SAR- and TIR-values for fish of the T1 and other transported smolts, all in comparison with the C0 and T0 cohorts. We need proper TIR-values for all of the sampled cohorts, measured Lower Granite Dam to Lower Granite Dam. The retrospective report reminds us that the Lower Granite Dam to natal source component of survival is outside the current CSS mandate, but suggests that we need to know more about it. We concur. (See Chapter 6 of the CSS Retrospective, designed to evaluate survival of hatchery and wild Chinook on their return journey from Bonneville to Lower Granite Dam (LGR) and from LGR to their source hatchery or spawning ground, and section V.E of this ISAB/ISRP review.)

IV. Evaluation of the Effectiveness of the CSS Retrospective in Answering Concerns Raised by the ISAB Review of the CSS 2005 Annual Report (ISAB 2006-3)

Each recommendation relating to an ISAB concern from the review of the CSS 2005 Annual Report is stated, and then followed by an assessment of how well the CSS Ten-year Retrospective Report responds to the concern.

> *1. The CSS needs to more effectively present the methodologies used in their analyses so the criticism of complicated and convoluted formulas can be avoided. The scattered explanations in several annual progress reports could be consolidated. Clear definitions of all notation used are required to avoid confusion caused by the use of the same notation for different concepts.*

Comment: The report provides a substantial improvement in the presentation and explanation of the sophisticated methodologies used in analyses of CSS data. The scope of CSS investigations resulted in an extensive report containing many detailed summaries of past and present work. The CSS team has responded very well in a short time frame to the difficult challenge of including enough details to allow scientific review without obfuscation by sheer volume of material.

> *2. The ISAB agrees with critics who express concern that the two downriver sites (Carson Hatchery and John Day River) are probably too few to give accurate upriver-downriver comparisons of SARs. This concern is bolstered by the variability among upriver hatcheries shown by the CSS data. For this upriver-downriver comparison to be generally accepted, it seems prudent to add more downriver sites in the future.*

Comment: In this report, the ISAB and ISRP no longer recommend adding additional downriver stocks for upriver-downriver comparisons. We understand this is a critical change in our recommendation and reach this conclusion after significant internal discussion based on findings from our latent mortality report and evaluation of the CSS Retrospective Report.

In the current analyses the limitation of sites continues to be a problem. The unbalanced upriver-downriver design makes tenuous the comparison of SARs among five Snake River Basin hatcheries compared to SARs at one downriver hatchery. Similarly, comparing one wild John Day population from one subbasin with wild Snake populations from three subbasins with geographical variation in habitat types, productivity, predator populations, and local climatic conditions makes interpretation problematic. However, this is not the core reason for recommending this analytical approach be discontinued. 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. Moreover, there is no parallel river system, so the challenge is more difficult than simply finding populations that could serve as replicates. 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. Finally, the absence of a clear measure indicating the cause for differences in upriver-downriver SARs does not mean that a real effect is absent.

The ISAB's Latent Mortality Report (ISAB 2007-1[4]) expresses in detail the concerns about interpretation of cause-effect from the upriver–downriver statistical comparison and notes also the lack of a needed appropriate baseline, the damless reference, in the context of attempts to measure latent mortality:

> *"Although the discussion concerning interpretation of stock and recruitment analyses, environmental covariates, and the value of upstream (Snake River) versus downstream (John Day River and Carson hatchery) continues (Hinrichsen pers comm, Paulsen and Fisher in review, Schaller and Petrosky in review), the ISAB questions whether this continuing discussion is productive.* (ISAB 2007-1, p.16)*"*
>
> *"One thing that is clear is that we are not able to estimate latent mortality for the damless reference condition.* (ISAB 2007-1, p.18)*"*

In the 2006 CSS review, the ISAB (ISAB 2006-3, pg. 12) recommended that more lower-river sites/stocks be included in the analysis. We now doubt that there are a sufficient number of appropriate downriver wild stocks available to make a meaningful comparison. The further downriver one goes in seeking more wild stocks, the more likely there will be climatic, watershed, habitat, and life history differences between the Snake River and lower river sites, thus exacerbating the physiographic and habitat differences between those sites. Finally, even if statistical differences between upstream and downstream stocks were found with increased sample size, it would be extremely difficult to determine the actual causes of the difference. 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.

In regard to life history differences between upriver and downriver stocks, the sponsors maintain that appropriate evaluation of these differences requires historical data contemporaneous with dam construction (p. 144). It is uncertain whether these data are available, but, lacking historical data, the sponsors compared life history characteristics using current data, thus violating their own supposition that the comparison should be made using data contemporaneous with dam construction. Perhaps the sponsors have addressed life history differences as well as possible given the available data, but the analysis of the current data does not shed much light on life history differences that might have been present at the time of dam construction.

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. The sponsors now appear to be engaged primarily in addressing the

[4] ISAB Latent Mortality Report: www.nwcouncil.org/library/isab/isab2007-1.htm

assumptions. They may have made their case as well as it can be made, and little may be achieved from additional analysis, due in part to the difficulty of validating the assumptions for the reasons given above.

Although not necessarily part of the CSS project, perhaps 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. Some of this monitoring and evaluation is likely already taking place outside the CSS project, but an analysis of those efforts was beyond the scope of this ISAB/ISRP review. In any event, the additional monitoring and evaluation of downstream stocks ***should not*** be directed toward upstream-downstream comparisons, because it is still unlikely that the problems given above could be adequately addressed.

3. *Data on size of all PIT-tagged fish from hatcheries and other release sites should be included in the report in much greater detail. Size at release may be a significant factor in differential SARs. The ISAB recommends including a specific section in the report focusing on the potential effects of size at release on survival of all PIT-tagged fish.*

Comment: Data on the size of PIT-tagged fish from hatcheries and other release sites have been included in the report. There were statistically significant differences in smolt size by release site. Although size at release may be a significant factor in differential SARs, the analyses presented in the retrospective report found no consistent or systematic difference in size-at-migration between up-river and downriver smolts. It should be noted that smolt size is recorded at time of tagging, not at time of release.

4. *Assumptions inherent in the analyses should be specifically evaluated, with continued vigilance toward avoiding bias in this evaluation.*

Comment: The simulation studies presented in Chapter 7 explore the impact of violations of the Cormack-Jolly-Seber (CJS) model assumption. The results suggest that the CJS model output is robust to all but the most extreme within season variation in survival or detection probability. No doubt other sensitivity analyses could be pursued, but the current study should provide substantial confidence in the CSS team's results. Other important assumptions, such as treating $SAR(C_0)$ as a binomial proportion for estimating sampling variance, have also been explored in the report.

5. *Pre-assigning the intended routes of passage at the time of release into in-river and transport groups would greatly simplify calculation of SARs and eliminate much criticism of current methods that are unnecessarily complex. This modification to the study design is scheduled for implementation in 2007.*

Comment: Pre-assigning has been done starting in 2007 and this should help reduce the complexity of the analysis protocols in future years.

6. *Analyses could emphasize more diverse metrics of differential survival, thus avoiding the criticism that the project staff focuses mainly on contentious issues such as the relative survival of transported and in-river migrants (T/C ratios) and differential delayed mortality between transported and in-river migrants (D). Passage routes, numbers of dams bypassed, distance from ocean, different hatchery practices, and other features have been explored beyond the issue of transportation.*

Comment: The report concentrates on SARs, TIRs, S_R, and D, and their traditional analyses. Almost nothing has been addressed related to this recommendation with the exception of instantaneous mortality, Z. 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. The hatchery tag recovery data could be used to provide comparative evaluation of various hatcheries and hatchery practices resulting in recommendations for hatchery management or production goals. TIR seems the best measure to address the issue of the relative survival of transported and in river fish, assuming that C_0 truly represents the appropriate control population and that jacks are not counted among the returns. The value of long data series collected and analyzed in the same way is clear. However, in this case it would be beneficial to also do parallel analyses using C_1s. We also suggest that analyses using jacks should be completed to resolve, or more fully understand, discrepancies with other studies.

7. *Often in summary reports of large data sets, there is a tendency to use a lot of references/citations to reduce extensive length of the document. The ISAB in its review of this report wants to see the key data and data summaries (supporting major conclusions) in the body of the report or in appendices included as part of the report. Having to search through numerous annual reports, other hard to find technical reports, or publications reduces the prospects of a timely review.*

Comment: The report presents key data and data summaries in support of their major conclusions. The body of the text is long for a summary report, reflecting the scope and complexity of the material summarized. Appendices could have been used even more to handle the extensive details.

V. Chapter by Chapter Specific Comments for the CSS Team

It should be noted that, because this is a retrospective report, much of what is presented is a summarization of earlier reports, without specifically referencing them. There have been many annual reports and a lesser number of publications that are too voluminous to reproduce here. In later chapters, some alternative statistical methods and approaches are introduced, incorporating previous suggestions of the ISRP, ISAB, and various other reviewing bodies. This development is encouraging, but it illuminates one of the limitations of a 10-year retrospective: we lose the intellectual time-track. The retrospective report does a better job of tracking management changes than recapping the intellectual path over the past decade. Embedding citations of earlier sources explicitly in the text would clarify the development of the methods.

The review comments that follow are intended for the CSS team's benefit to better understand the ISAB's perspective on the occasion of this ten-year retrospective activity.

A. Chapter 2: Travel Time, Survival, and Instantaneous Mortality Rates of Yearling Chinook and Steelhead through the Lower Snake and Columbia Rivers, and their Associations with Environmental Variables

This chapter presents methodology and results for estimation of travel time, survival, and a measure of instantaneous mortality for the years 1998-2006 for two multi-project reaches, Lower Granite Dam to McNary Dam (LGR-MCN) and McNary Dam to Bonneville Dam (MCN-BON). Data for hatchery/wild yearling Chinook salmon and steelhead are analyzed. Within year trends were monitored by identifying eight weekly cohorts within each season. The relationships between estimated travel time, survival, and the instantaneous mortality measure, as well as several environmental covariates were explored with log-linear multiple regression.

Overall, Chapter 2 is a major improvement on individual annual reports on this topic. It is a worthwhile step in synthesizing the results to date. It is also written and presented very well in most places, and the organization and logic are not difficult to follow.

Much of the material in this chapter has been previously released and often evaluated. An exception is the estimation and interpretation of Z, the proposed measure of instantaneous mortality rate. The use of *Z* has its attractions, because any assessment of mortality over a short time period (as through a bypass, through an RSW, through a project, or down a short portion of a single reach, etc.) needs something approximating an instantaneous rate. There has been steady pressure from reviewers of the survival data for *shorter* time-step assessment of mortality, relative to variables such as alternative hydrosystem options, alternative routes of passage, temporal variation, and variation in different stocks.

The quantity Z, is clearly related to survival because for any short time period, say Δt, $S = \exp\{-Z \bullet \Delta t\}$. It seems clear from what is presented that the statistical behavior of *Z* is better than that of *S*; the distribution is more nearly symmetric, and the variance is substantially more homogeneous over the range of survival rates of interest. For assessment of the impact of virtually all environmental, biological, and managerial variables on survival, the standard strategy is to regress $-\log S / \Delta t = Z$ on those variables. Thus, *Z* is not really new, but the real issue is how well an instantaneous rate can be estimated, in view of the way survival data are gathered. Estimation limitations for *S* translate into derivative limitations for *Z*, among them:

(a) The numbers of tagged fish released in a given year are large, but if the time interval used for *instantaneous assessment*, Δt, is too small, the estimation noise is unmanageably large. Some aggregation of data is necessary to obtain credible sample sizes, a constant challenge for survival analysis. Typically, that aggregation translates into an extended Δt, for example pooling one to three weeks, a third of a season, through a reach, sometimes through all upriver projects or through all downriver projects. Depending on the sample sizes needed for a particular problem, there is substantial averaging over time. The end result is that the analysis is typically using $\overline{Z}$ over a time period of some length, not *instantaneous Z*.

(b) Cormack-Jolly-Seber (CJS) methodology tallies fish passing more than one project, in estimating *S* (hence *Z*). That means smolts spend time in the river and in the impoundment behind a dam, probably not homogeneous survival environments. The term *instantaneous mortality* rate is an *approximate* choice of descriptor, though the choice of *Z* is better than using untransformed *S* for analysis, with all due allowance for their joint limitations.

From the modeling point of view, instantaneous mortality (*Z*) is most useful if it is applicable at most (or at least some) known instants. As defined and estimated in the CSS report, Z does not disentangle dam passage from traversing the reach between dams. As a result, it may not yet be applicable at any instant. Instantaneous mortality rates may be very different during dam passage compared to migrating through a reservoir. Good estimates of *Z* for the traversal of each reservoir would be quite useful and might provide better estimates of the probability of survival for the various passage routes through the dams.

In reviewing earlier COMPASS (ISAB 2006-2) and Latent Mortality (ISAB 2007-1) reports, the ISAB came to the conclusion that log S (equivalently, $\overline{Z}$) is not reliably linear in predictor variables over any extended range of S, though for short ranges, the approximation is fairly reasonable. We have repeatedly suggested that logistic-linear models would provide better statistical behavior, and might yield less complex predictive models, requiring fewer interaction and polynomial terms than those previously reported. As in earlier reviews, the ISAB's encourages exploring a wider array of statistical strategies in CSS context. For later chapters, presumably in response to earlier reviews, the CSS team has switched to logistic treatment.

We note that, among the environmental covariates used as predictors, there are cases where Water Travel Time (WTT) is the better predictor, but other cases where $(\text{flow})^{-1}$ is the better predictor, though the difference in predictions is usually quite subtle. The predictions should be close, of course, because the two variables are highly correlated. Both seem to be better generic predictors than (flow) itself. Our sense is that one could choose one or the other for general usage, with no meaningful loss of performance in any particular case. The report mentions that there are threshold effects, beyond which there is no change in survival, and we wonder whether some sort of hyperbolic treatment of either $(\text{flow})^{-1}$ or WTT would be worth exploring, rather than fitting splines, quadratics, and/or interaction terms whenever linearity fails.

Our other generic comment, echoing some of the agency reviews, is that the array of multiple regression analyses described may be providing little marginal benefit. For example, the exploration of several different weighting schemes seems unnecessary when careful theoretical considerations should provide some *a priori* guidance. The ISAB and ISRP realize that the CSS team is making an extensive effort to discover effective methodologies and provide answers under a variety of scenarios, but we recommend the strategy of using simple models whenever possible.

B. Chapter 3: Annual SAR by Study Category, *TIR*, S_R, and *D* for Hatchery and Wild Spring/Summer Chinook Salmon and Steelhead: Patterns and Significance

Chapter 3 includes an extensive discussion of study methods, along with the techniques and formulas that were used. The report provides a summary of results through the life of the study and an overall comparison of study results among years. It provides clear interpretations of these results and summary conclusions. This chapter is devoted to a comparison of in-river and transported fish, where the sampling unit of interest is the entire yearly run. The chapter describes the computation of SARs for three cohorts, T_0 = fish transported from Snake River Dams at Dams at Lower Granite (LGR), Little Goose (LGS) and Lower Monumental (LMN), C_0 = in-river fish not detected at any of the projects, and C_1 = fish detected at LGR, and then diverted to the bypass system, but otherwise migrating in-river. Tagged fish were, in all cases, assumed to mimic the behavior patterns and exhibit the survival experience of their untagged counterparts. Not everyone will agree with all of the conclusions, but the process and evidence used to arrive at those conclusions are transparent, so the argument becomes one of interpretation.

Chapter 3 and appropriate appendices clearly lay out the methods, procedures, formulae, and data used to establish the reported values of SARs, TIRs, S_R, and D. The report summarizes the experience for spring-summer Chinook, both hatchery and wild stocks, and for Steelhead, both hatchery and wild stocks. The net benefit of transportation would appear to decrease in the order (hatchery and wild Steelhead) > (hatchery Chinook) > (wild Chinook), with wild Chinook showing no consistent evidence of any net benefit, $TIR_0 \sim 1$. Even where there was substantial benefit from transportation, none of these SARs were sufficient for stock persistence, so while it sometimes does improve survival, transportation alone is not sufficient to ensure persistence. The broad outline of results seems clear enough, but we must emphasize some of the limitations pointed out by the CSS team, and flag some others that will need attention in the future.

1. The assumption that tagged fish are typical of the untagged fish in their respective cohorts is difficult to test in any empirical fashion and is a bit worrisome, in view of two features flagged by the CSS team. First, in the early years of tagging for both species, untagged fish arriving at the collection facility at LGR were routinely transported (joining the T_0 cohort), but tagged fish were diverted to the bypass system (joining the C_1 cohort). 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. In more recent years, a substantial fraction of tagged fish has been added to the T_0 cohort, so the discrepancy is less pronounced. Second, for both species, both hatchery and wild stocks exhibited subsequent $SAR(C_0) > SAR(C_1)$. The CSS report attributes the extra losses in the C_1 cohort to handling at the LGR collection facility, associated with diversion to the bypass system. Whether that extra survival toll on tagged in-river smolts is due to handling at LGR or to tagging itself, in-river smolts would appear to suffer from the process, relative to untagged T_0 fish.

2. Transported fish are collected at three projects (LGR, LGS, and LMN), but are all treated as though released at LGR. In order to do this, the survival experience of fish collected at LGS and LMN is "back-translated" to LGR, using the LGR_LGS and LGS_LMN reach survival probabilities. Those probabilities are very imprecisely estimated, as pointed out by the CSS team, and that imprecision translates into larger variances for S_T, and from there into low precision for SAR(T_0). The wild Chinook and Steelhead sample sizes are small enough to be problematic in any case, so inflating the error of S_T is unfortunate. 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.

3. 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.

4. In an extremely low flow year, such as 2001, the C_0 cohort was not available, so TIR was defined as SAR(T_0)/SAR(C_1). SAR(C_1) is an inappropriate reference point and the returns of C_1 were so few that the estimate itself is shaky. Although this is the only estimate available, this TIR estimate is probably severely inflated. The CSS team is to be commended for finding creative alternatives for estimating TIR in 2001, but regardless of the TIR estimate, transportation may be the only strategy available in a severe drought, given current circumstances.

5. The report clearly describes the limitations imposed by small sample sizes for wild Chinook and for Steelhead in estimating TIR_0. Sample sizes of hatchery Chinook are large enough that the evidence for $TIR_0 > 1$ was collectively convincing. Transportation of steelhead seems beneficial enough that even limited precision is not a serious impediment to establishing that fact. For wild Chinook, also represented by inadequate sample sizes, the noise is larger than the signal for TIR_0, and nothing

6. g very convincing can be determined.

C. Chapter 4: Estimating Environmental Stochasticity in SARs, *TIR*s, and *D*s

Chapter 4 produces statistical distributions of SARs, TIRs, and Ds over multiple migration years with the goal of estimating environmental stochasticity for these metrics. These distributions are produced for each collector project with the purpose of allowing prospective modeling and for monitoring and evaluation to estimate variation in observed TIRs. The work presented represents only category C_0 fish as the in-river population but the CSS team notes that the methodology can be extended to other in-river populations, presumably category C_1 fish.

The approach is based on the assumption that long-term distributions of SARs can be modeled as a beta distribution. This flexible distribution is often chosen as a model for data restricted to values between 0 and 1. The material presented here extends earlier work (Berggren et al. 2005[5]) to wild steelhead, to estimate SARs, TIRs, and Ds for each transport project separately, and to modify the method for TIR distributions that include covariance between transport and control SARs, by using beta random variables.

An estimate of environmental variance of SAR is obtained by removing an estimate of demographic variance from total inter-annual variance. Most of the analyses focus on annual SAR estimates with their inherent limitations of utility, as noted elsewhere in this review. The CSS team observes that the migration season could be broken into segments, based on arrival timing and the method used to examine within season patterns. The ISAB and ISRP encourage this additional investigation to provide value for project management.

As in any study involving complex adaptation of statistical methodology, there will be differences of opinion about details. Without dwelling on these details, we note that there may be bias in the estimation of the sampling variance that results in a derivative bias in the estimate of the environmental variance. The CSS team is in a good position to investigate potential biases, and they do identify that there is a slight underestimation of environmental variance in an analysis where there is a positive correlation between smolt number and adult returns. Support for some assumptions, such as binomially distributed sampling error in SAR estimates, has been provided.

[5] Berggren, T., H. Franzoni, L. Basham, P. Wilson, H. Schaller, C. Petrosky, E. Weber, and R. Boyce. December 2005. Comparative Survival Study (CSS) of PIT-tagged Spring/Summer Chinook and PIT-tagged Summer Steelhead, 2005 Annual Report, Mark/Recapture Activities and Bootstrap Analysis. BPA Contract # 19960200. 107 pages. Available at www.fpc.org/

D. Chapter 5: Evaluation and Comparison of Overall SARs

Chapter 5 presents a wealth of original study data and analyses to try and evaluate: (1) how well overall SARs for Snake River listed salmon and steelhead are meeting the NPCC (2003[6]) interim biological recovery goals, (2) how certain environmental factors influence the overall SARs, and (3) compare SARs for Snake River spring/summer Chinook and SARs from downriver populations to relate to effects of the hydrosystem in the lower Snake River.

(1) Overall SARs for wild spring/summer Chinook (geometric mean 0.9%, range 0.3-2.4%) and wild steelhead (geometric mean 1.6%, range 0.3% - 2.9%) fell short of the NPCC SAR recovery objectives (2% minimum, 4% average). These conclusions were supported by the data presented in the report, but do not include the untagged fish data. For some years, the untagged fish have been estimated to survive at considerably higher rates than tagged fish (Copeland et al. 2007[7]). Regardless, these SARs are low and are reason for concern for recovery of these listed stocks. We note that the interim NPCC recovery goals may be revised by the final recommendations of the Technical Recovery Teams, soon to be available, and future recovery goals will most likely be tailored to individual ESUs.

(2) CSS analyses indicated that SARs of Snake River wild spring/summer Chinook were positively correlated with faster water travel time (WTT) experienced during the smolt migration, cooler phases of the PDO index (primarily in May or September), and stronger down-welling in the fall (November) during the first year of ocean residence. SAR estimates based on a limited time series of PIT tag data have a high degree of uncertainty, and environmental and ocean conditions vary on several temporal and spatial scales.

(3) In recent CSS reviews, both the ISAB and ISRP have expressed concern about the limited number of sites and stocks used in the upriver/downriver comparisons. Conclusions regarding the effects of the hydrosystem that are based on upriver/downriver comparisons are not scientifically defensible. The observation that SARs for wild Chinook were only one quarter that of similar downriver populations that migrated through fewer dams does not provide proof that the hydrosystem is responsible for this SAR difference; there are myriad alternative possible explanations. The comparisons are weak because of the limited number of sites compared and the multitude of confounding variables that could influence SARs, in addition to the hydrosystem effect.

Although both the ISRP (ISRP 2006-6) and the ISAB (ISAB 2006-3) (see above) have recommended that more lower river sites/stocks be included in the comparative analysis, we now doubt that there are a sufficient number of downriver wild and hatchery stocks available with adequate quantity and quality of data to provide a meaningful comparison. It is very likely that there will be climatic, watershed, and habitat differences between the Snake River and lower

[6] 2003 Mainstem Amendments to the Columbia River Basin Fish and Wildlife Program: www.nwcouncil.org/library/2003/2003-11.htm

[7] Copeland, T., J. Johnson, and G. Bunn. 2007. Idaho natural production monitoring and evaluation. Report of Research by Idaho Department of Fish and Game to Bonneville Power Administration, 2005-2006 Annual Report, Project No. 199107300 (BPA Report DOE/BP-00023363-1), 53 electronic p.

Columbia River sites, thus exacerbating the physiographic and habitat differences between those sites. Even if statistical differences in SARs between upstream and downstream stocks were found with increased sample size, it would be extremely difficult to determine the causes of the difference.

While the CSS team has addressed life history differences as well as possible, given the available data, the current data shed no light on life history differences that might have been present at the time of dam construction. Too many (salmon) generations have elapsed since then, and the evolutionary pressures to adapt to a new situation have been severe. Recent evidence from outside of the Columbia River argues that salmon life-history traits can evolve rapidly. Extending these observations to the Columbia River, the salmon have likely adapted and evolved (as well as they can) in the time elapsed. The Snake River stocks are, in all likelihood, no longer the same as their pre-hydrosystem ancestors. The downriver stocks have also been under severe pressures from their own changing environment, and they have also evolved, but almost surely not in the same direction as the Snake River counterparts.

E. Chapter 6: Partitioning Survival Rates - Hatchery Release to Return

This chapter is designed to evaluate survival performance of Chinook (both hatchery and wild origin) on their return journey, (1) the BON-LGR leg, and (b) from LGR to their source hatchery (or spawning ground). Steelhead will be evaluated separately and in the future. Three cohorts are being examined here for each stock, T_0 (transported from LGR), T_1 (transported from LGS or a lower project), and C_0 (in-river). We commented in reviewing Chapter 3 that it would be useful to separate T_0 from T_1 fish, and the results of Chapter 6 make the benefits of that recommendation clear. The central messages emerging from contrasting T_0 and C_0 are as follows:

<u>BON → LGR</u>

a. 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. $SAR(C_0)$ is also a bit better than $SAR(T_1)$, but differences are quite a bit smaller and are generally non-significant.

b. Logistic regression modeling shows that all attractive models include a term for a transportation effect, whether or not environmental or management effects are added, and that sometimes, terms for variables other than transportation are not particularly helpful. It is worth recalling that $TIR_0 = SAR(T_0)/SAR(C_0) > 1$ for hatchery Chinook and ~ 1 for wild Chinook, measured from LGR-LGR (Chapter 3).

<u>LGR → spawning ground</u>

c. There are serious detection issues above LGR, due to the number of other factors that can obscure detection at the natal spawning sites, among them harvest or straying, but about

60% of the migrating adults that pass LGR are subsequently detected at a spawning site. Upward adjustment for those that are not detected yields an estimate of about 75 - 80%, depending on the cohort, but the CSS report states that the estimates are not very reliable, given the limited auxiliary information available.

d. Based on preliminary analyses of incomplete data, there is suggestive evidence that the farther upriver that natal source is from LGR, the greater the probability of returning to it, interpreted in terms of the time for imprinting on the way downstream from source to LGR.

Four larger messages emerge from Chapter 6.

1. The retrospective report reminds us that LGR → natal source component of the survival experience is outside the current CSS mandate, but suggests that we need to know a lot more about it and that our evaluation needs to improve. We concur.
2. Transportation downriver, whatever benefits it may have on the outward journey, has substantial negative consequences on the homeward journey.
3. 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 T_1, T_2, and so on have been "back-adjusted to LGR" and are lumped with LGR-transported fish (T_0). That is not necessary, and indeed, it is time for a direct assessment of SAR- and TIR-values for fish of the T_1 (LGS-transported), T_2 (LMN-transported), and other transported smolts as well, all in comparison with the C_0 and T_0 cohorts.
4. Proper TIR-values are needed for all of these cohorts, measured LGR-LGR, but the data exist now, and the possibility of an improved transportation strategy is worth evaluating with the already available information.

F. Chapter 7: Simulation Studies to Explore Impact of CJS Model Assumption Violations on Parameter Estimation

Chapter 7 goes a long way toward establishing the robustness of the Cormack-Jolly-Seber (CJS) results. The simulation studies presented in Chapter 7 explore the impact of violations of the CJS model assumption. In particular the CJS model assumes no variation in detection and survival probabilities through the season. Within season variation in migration conditions could cause variation in either detection or survival probabilities, violating the CJS assumptions. The CSS group has executed extensive simulation studies of the sensitivity of their procedures to variation in these probabilities through a season. Their approach is a standard method for establishing these sensitivities, and their results suggest that the CJS model output is robust to all but the most extreme within-season variation in survival or detection probability. Other sensitivity analyses could undoubtedly be pursued, but the analyses presented here provide substantial confidence in their current CJS results.

G. Chapter 8: Conclusions and General Recommendations for Future Direction

Chapter 8 provides a useful overview of the accomplishments and chapter-specific conclusions of the 10-year retrospective report. The chapter also includes overall conclusions and suggestions for future activities from the CSS team.

The overview notes that CSS constitutes a successful implementation of a large-scale tagging program, and we agree. PIT-tags from other marking programs have been used to augment funded CSS tagging. We encourage even more cooperation among PIT-tag marking programs to address critical uncertainties and improve the reliability of survival estimates.

Regression models for the LGR-MCN and MCN-BON reaches for both Chinook and steelhead that relate fish travel times, survival, and instantaneous mortality rates to environmental variables, such as water travel time, average percent spill, flow, and Julian day, result in an extensive array of relationships. It would appear that a concerted effort at developing more parsimonious models is necessary. Models that differ by species and reach should generate more hypotheses for future investigation.

As noted elsewhere in this review, we urge caution in interpreting results and suggest that assumptions incorporated in making interpretations be clearly presented and measures of uncertainty always be included. As a simple example when stating that transportation provided, or did not provide, benefit to a species or wild/hatchery group the possibility of selection bias of fish for transportation due to variables such as size, condition, and location in the water column should be part of the interpretation. We also note that conclusions about mortality or delayed mortality of transported fish relative to in-river fish are not equivalent to saying mortality or delayed mortality DUE to transportation, unless all other factors are accounted for. Similarly, statements that trends are consistent with a specific hypothesis are most useful if alternative explanations are examined and discounted.

The overall self-assessment at the end of the chapter is useful. The four primary objectives stated are: 1) develop long term indices of transport and in-river SARs for Snake River hatchery and wild spring/summer Chinook and steelhead; 2) develop long term indices of survival rates from release of yearling Chinook smolts at hatcheries to return of adults at hatchery; 3) compute and compare overall SARs for selected upriver and downriver spring/summer Chinook hatchery and wild stocks; and 4) begin a time series of SARs for use in regional long-term monitoring and evaluation. We agree that long-term indices of transport and in-river SARs have been developed (objective 1) and that a time series of SARs for use in monitoring and evaluation has been initiated. As noted by the CSS team, improvement in estimates of adult returns to hatcheries is necessary to fully meet objective 2. As stated elsewhere in this review, the interpretation of the comparison of upriver/downriver stocks is problematic. This necessitates a re-evaluation of the purpose of this objective, as noted previously.