FISH PASSAGE CENTER

1827 NE 44th Ave., Suite 240, Portland, OR 97213
Phone: (503) 230-4099 Fax: (503) 230-7559
http://www.fpc.org
e-mail us at fpcstaff@fpc.org

August 31, 2007

Robert J Austin
Deputy Director of Fish and Wildlife
Bonneville Power Administration
PO Box 3621
Portland, Oregon 97208-3621

Dear Mr. Austin:

Thank you for your review of the Draft, Ten Year Retrospective Summary Report. The following response was developed by the Comparative Survival Study Oversight Committee, (Committee) comprised of, the Columbia River Inter-tribal Fish Commission, the Washington Department of Fish and Wildlife, the Oregon Department of Fish and Wildlife, the Idaho Department of Fish and Game, and the US Fish and Wildlife Service. As you are aware the Comparative Survival Study is a joint project of the agencies and tribes. The study design, the implementation of the study and the analysis are carried out collaboratively among the sponsoring fish and wildlife management agencies. The Committee has developed the following response to your general comments, followed by the response to each specific comment.

General Comments
As with past BPA review comments, we found several comments which will be helpful in improving the overall strength of the final report. However, many of the BPA general comments summarized in the cover letter are presented in such general terms without an explicit context that they are difficult to address. They are presented as sweeping conclusions of a critical nature without any basis provided. Further, some of the general statements are inaccurate and some of the reviewers' specific comments are erroneous.

Transparency, reproducibility, data, detailed methods, tagging results
A majority of the BPA conclusion comments addresses the issues of transparency, reproducibility, data and detailed methods. BPA states that the study, methods and data are not reproducible. We disagree with the BPA statements.

All of the data, detailed methods and mathematical derivations are available. The attached (attachment 1) email documents that on June 12, the FPC received a request from staff of Jones & Stokes, reviewing the Ten Year Retrospective Report under contract with BPA. On June 13 the FPC, in response to this request, transmitted 61 files, providing the specific capture history

input files for each of the 2,413,209 fish included in the ten year report. In addition, in the email response, we indicated that FPC staff are available to answer additional questions to assist the consultants' work. With the input files and the formulas, BPA and or their consultants should have been able to generate the components for the formulas, using the widely available MARK or SURPH programs, and then use those components in the formulas in Appendix B of the draft report, or methods explained in the report chapters. In any case the CSS Oversight Committee and the FPC were available to assist reviewers as indicated in the attached emails.

The BPA comment does not explain how BPA and/or their consultants tried to reproduce results. Consequently, it is difficult to respond to the BPA comment regarding reproducibility. Neither BPA nor their consultants attempted to contact the FPC or the Oversight Committee with questions or requests for additional information. BPA and their consultants neither requested a meeting to discuss their attempts to reproduce results nor explained in their comments what specific attempts they made to reproduce results. As always the CSS Oversight Committee and the FPC are available to discuss the report with BPA and their consultants. All of the specific data and the mathematical formulas have been provided to BPA and/or their consultants, and our willingness to respond to additional questions was indicated. Given this lack of information on what more BPA feels they need, we can't determine how to address BPA's request for additional "transparency".

Missing information

BPA states that information is missing and specifically states that formulas for calculating SARs are missing. This is inaccurate. Specifically, Appendix B of the Ten- year Report includes all of the mathematical derivations for the formulas utilized in the Chapter 3 analysis; these include the formulas for calculating SARs. In addition, Chapter 3 includes the formulas for SARs.

Non-standard modeling practices

We disagree with BPA's contention regarding non-standard modeling practices. We have utilized generally accepted, standard statistical procedures for estimation, model-building and associated analyses. Analyses new to this report are based on extensions to methods developed in referenced peer reviewed literature, and methods and assumptions are clearly spelled out. The CSS ten year report is being peer reviewed in this process and CSS products have been peer reviewed in previous years,

Inability to reproduce results

BPA or their consultants' inability to reproduce results do not reflect on the scientific rigor or analytical procedures, modeling or methods used in the Report but perhaps problems with BPA and or consultants attempts. BPA has not described the process used to attempt to reproduce the CSS results, nor did they describe what specifically they were trying to reproduce. They have not availed themselves of the offer by the CSS Oversight Committee to provide guidance or answer questions. All of the input files were available to them and all of the mathematical derivations and formulas were provided in the report.

Latent Mortality

We found a difference in instantaneous mortality rates between SARs of Snake River and downriver wild spring/summer Chinook populations, similar in magnitude to that estimated

previously in published literature from spawner-recruit data. The level of differential mortality was relatively small only between upriver and downriver hatchery Chinook (as stated in BPA comments). The BPA or their consultants' proposed adjustment to differential mortality has two major flaws. The BPA adjustment is inconsistent with the definition of differential mortality, and it fails to account for passage survival of transported smolts.

Tagging Results

All of the tagging files and the individual capture history records for each fish were provided to BPA consultants as previously documented. All of the resources of the "CSS organization" were offered to BPA consultants. BPA and consultants did not make any contact with the CSS Oversight Committee or technical staff in their undefined attempts to reproduce results.

Upstream downstream comparison

The BPA comment that data do not support an upriver/downriver comparison is not accurate. Differential mortality is estimable from both PIT-tag and spawner-recruit data. The ISAB (2006) recommended incorporation of additional downriver wild and hatchery populations into the comparison. The CSS Oversight Committee concurs with the ISAB recommendation, and has proposed, but not received BPA funding, to PIT-tag additional downriver populations.

Invalid assumptions

The BPA comment is inaccurate, and the proposed adjustment for in-river migration mortality is inconsistent with the published definition of differential mortality.

Detailed responses to each of the individual comments submitted by BPA are attached (attachment 2).

The CSS Oversight Committee is grateful for the significant investment by BPA in the review and preparation of comments on the draft report. The report has been improved as a result of addressing and incorporating comments. We look forward to future positive collaboration with BPA on future CSS monitoring and evaluation.

Signed

Michele DeHart

Michele DeHart
Project Leader, Comparative Survival Study

Attachment 1

From: Tom Berggren
Sent: Wednesday, June 13, 2007 2:13 PM
To: Kevin Malone
Cc: Howard Schaller (howard_schaller@fws.gov); Paul Wilson (Paul_H_Wilson@fws.gov); Steve Haeseker (steve_haeseker@fws.gov); Charlie Petrosky (cpetrosky@idfg.idaho.gov); Eric Tinus (eric.tinus@state.or.us); Tim Dalton (Tim.Dalton@state.or.us); Rod Woodin (woodirmw@dfw.wa.gov); Michele DeHart
Subject: RE: CSS Database

Kevin Malone:

Attached is a link to FPC's website from where you may download detection history data used in the CSS. There are 14 directories containing a total of 116 separate data files, which include all wild and hatchery Chinook and steelhead analyzed for SARs in the CSS 10-yr Retrospective Report. This data will be temporarily held on this site until the close of business on June 29, 2007, giving you 12 business days to access and download those data of interest to your review of our draft report. If you have any questions regarding file contents or field names, you may contact me by email.

Tom Berggren

Cc: CSS Oversight Committee members

The CSS file download webpage is at the following link.

http://www.fpc.org/css/css_files.html

From: Kevin Malone [mailto:kmmalone@wavecable.com]
Sent: Tuesday, June 12, 2007 8:08 PM
To: tberggren@fpc.org
Cc: mfilardo@fpc.org
Subject: This is Spam CSS Database

Hi!

I am reviewing the CSS report and would like to get the detection history database for this data set. Specifically, the juvenile and adult detection history for each PIT-Tag used to generate the SAR data etc.

You can send it via e-mail as a zip file or if you point me to a FTP site that would be great!

Thanks!

Kevin Malone
Jones and Stokes

No virus found in this incoming message.
Checked by AVG Free Edition.
Version: 7.5.472 / Virus Database: 269.8.15/847 - Release Date: 6/12/2007 9:42 PM

Attachment 2
General Comments

Another aspect of the report used parametric models to partition total variance of metrics into natural variation and measurement error. However, the assumption, for example, that SARs are binomially distributed is inconsistent with the mark-recapture models used to estimate the values. Underestimating sampling error will positively bias estimates of natural variation. The report needs to use goodness-of-fit tests to assess the model assumptions and compare their parameter estimates with those of the nonparametric variance component formulas provided. Their inferences concerning natural variation do not take into account their own findings on ambient effects, the historical distribution of those factors, or how influences such as global climate change many affect projections in the future.

Response: See responses to specific comments on Chapter 4.

Additionally, the CSS incorrectly claims to have addressed the question of whether smolt transportation compensates for effects of the FCRPS on survival of Snake River Chinook and steelhead. At most, the comparison of the SARs of transport fish and inriver fish indicates whether transportation is a viable management option; it is not equivalent to comparing transportation to migration through the unimpounded river. The question of the effect of the FCRPS on salmonid migration and survival is important. However, it is not addressed by the analyses presented in this report.

Response: A major goal of CSS, estimation of the efficacy of transportation, will be described more explicitly in the revised report. In brief, both the absolute realized SARs under the current system, and the ratio of transport SARs to in-river SARs are estimated. Combined with information derived from other sources, it's possible to gain insights on the effect the hydrosystem has had on life-cycle survival rates. It's true that comparison of life-cycle survival under transportation to migration through the unimpounded river cannot be made using information derived only from CSS. However, key components of the comparison include a parameter reflecting any delayed mortality due to transportation (*D*), recent in-river survival rates, and estimates of the proportion of fish transported under recent conditions. These parameters are estimated in CSS and these parameters have been used in models to compare different strategies, including a "dam breach" or "natural river" option (e.g. Peters and Marmorek 2001; Wilson 2003; Zabel et al. *in press*).

Chapter 2

$\hat{Z} = \frac{-\log(\hat{S})}{FTT}$ *(Eqn. 3) provides a convenient but biased estimate of the instantaneous morality rate. Properly, the maximum likelihood estimate of Z would be based on the likelihood model* $L = \prod_{i=1}^{N} Ze^{-Zt_i}$ *and estimator* $\hat{Z} = \frac{1}{\bar{t}}$ *[Eqn. 4] where* t_i *= lifetime for the ith fish* $(i = 1, K, N)$. *However, PIT-tag data do not provide lifetimes for the fish, only travel times for the survivors. Therefore, the PIT-tag data are incapable of estimating instantaneous mortality rates. Any*

relationship between the true estimates of Z [Eq. (4)] and that used in the report [Eq. (3)] may be appropriate at best and seriously biased as worst.

Response: Our estimates of *Z* (Eqn. 3, above) are the maximum likelihood estimates for *Z* (Seber 1982:216). Contrary to this comment, PIT-tag data not only provide data on the travel times of surviving fish, they also provide survival rate estimates for release cohorts through the CJS methodology that can be used to estimate *Z*. We agree that the estimator suggested (Eqn. 4, above) cannot be used to estimate *Z*, and find this comment to be a useless suggestion in this application. We agree that our use of Eqn. 3 is appropriate, but disagree that this maximum likelihood estimator of *Z* (Seber 1982:216) is seriously biased.

The report seemingly takes a shotgun approach to the analysis. In the results section, which weighing scheme and why its selection was not revealed. The weight selection should be objective.

Response: We used standard statistical methods in the analyses and objective criteria for model building and variable selection. The weighting scheme was objectively determined by the scheme that maximized the adjusted R^2 values for the predictions on the arithmetic scale. The weighting schemes chosen are provided in the tables describing the models evaluated.

Proper weighting should be inversely proportional to the variance except when the variance estimates is correlated with the response variable. In this case, the weight should be inversely proportional to the variance but adjusted to eliminate the correlation.

Response: We evaluated this suggested weighting approach, but found that it resulted in lower adjusted R^2 values than the other weighting methods we investigated.

The report states, "we examined the sign of the parameter coefficients for plausibility and eliminated models with implausible sign." This is a dangerous and potentially misguided approach to modeling. First, such an approach eliminates the possibility that new insights might be developed and assumes all preconceptions are correct. Secondly, it is unwise to directly interpret the sign (+ or −) of partial regression coefficients (Neter et al. 1996:290-291). Such signs do not necessarily indicate a positive or negative relationship between dependent and independent variables but, instead, adjustments of the model in the presence of other covariates. This unorthodox model strategy can lead to odd modeling results (see comments below).

Response: We eliminated the approach of examining parameter signs and now report model fit statistics for all models that were evaluated. As a matter of clarification, this section of Neter et al. (1996:290-291) is primarily focused on the effects of multicollinearity and does not indicate that it is unwise to interpret the sign of multiple regression coefficients.

The report states, "models were fit and ranked according to their AICC and BIC scores." However, many tables (e.g., Tables 2.7-2.11, 2.13) report AIC scores while other tables (e.g., Tables 2.12, 2.15-2.16) report AIC and AICC scores. What was actually done and reported needs to be clarified. For example, are the AIC values in Tables 2.7-2.11 actually AICC and "AIC" is a typo?

Response: All tables with model fit statistics provide the AICc, BIC, R^2, adjusted R^2, delta AICc and Akaike weights (w_i) for each model evaluated.

"Integrated models of fish travel time and instantaneous mortality, with each component modeled being a function of environmental covariates" are mentioned but never described. If a multivariate computational model was actually used, it needs to be provided, along with associated assumptions (providing Eq. 2.2 is inadequate).

Response: We provide equation forms, model fit statistics, and parameter coefficients for the models characterizing median FTT, Z, and S.

Julian day was found in several instances to help describe regression relationships. The implication of this covariate in the models must be described for it is unlike the other covariates considered (e.g., WTT, percent spill, etc.). Julian date is a surrogate for numerous factors that may have a within-season trend including smoltification, flows, temperature, turbidity, etc. If the purpose of the regression analyses is to describe environmental and hydrosystem factors affecting fish response, inclusion of Julian data obscures the results. In some instances, (e.g., Table 2.15-2.16), it does a very good job all by itself!

Response: We provide a description of possible seasonal effects that the Julian day covariate may be capturing. The use of Julian day as a covariate to capture seasonal effects is a common modeling strategy with these data (Berggren and Filardo 1993, Smith et al. 2002, Williams et al. 2005). However, these possible effects (smoltification, photoperiod, fish length/size, predator abundance/activity) are those which are *not* already captured by the other variables examined (flows, temperature, turbidity).

AIC scores cannot be compared across different data sets (Burnham and Anderson 2002:80-81). Comparison of models of FTT and instantaneous mortality versus direct survival is inappropriate and Table 2.2 should be eliminated from the report.

Response: We used the same data set (observed survival rates) to compare with the predicted survival rates (predicted using three different approaches) using AIC values (Burnham and Anderson 2002:63). The table referred to has been expanded.

The authors are totally misinterpreting their estimates of instantaneous mortality Z. In this paragraph, they are equating Z to probability of mortality which is wrong.

Response: For values of $Z \leq 0.1$, mortality rates and Z estimates are approximately equivalent (Ricker 1975). However, to clear up any confusion on the trivial differences between the two, we have provided both daily percent mortality estimates and Z estimates.

The symbolism for box and whisker plots is not universally consistent or known. Captions should explain the symbolism.

Response: Box and whisker plots have a consistent definition and are an elementary topic commonly covered in rudimentary statistical methods courses. The first box and whisker plot now contains a description of what a box and whisker plot represents, for those who are unfamiliar with basic statistical concepts and data descriptions.

Caption fails to indicate which models the results refer to.

Response: The caption now indicates that the survival predictions are based on the variable *Z* approach described.

Omit because AIC are not comparable across different datasets.

Response: The same data set (observed survival rates) was used to judge the different approaches for predicting survival rates using AIC values. In addition, we used root mean squared error, R^2 values and the number of estimated parameters to judge the accuracy of the different survival modeling approaches.

Captions are inadequately described. Symbols for models are cryptic and need to be explained for clarity of interpretation.

Response: Table captions now provide a full description of the symbolism for the variables examined.

The selection of models examined is at times eccentric: Models may include an interaction term without one or both of the main effects included. Purpose of an interaction term is to modify the main effects; it is unclear what the interaction term means in the absence of the main effects.

Response: Models with an interaction term now include both main effects, even though better fits were obtained by omitting one of the main effects in some cases, as was shown previously.

The selection of models examined is at times eccentric: Higher-order polynomial terms are included in models without corresponding lower-order terms, which is not conventional in linear models; for example, squared term without the linear term.

Response: Models with second-order terms now include single-order terms, even though better fits were obtained by omitting single-order terms in some cases, as was shown previously.

Wonder whether this nonconventional approach to modeling is a direct consequence of dropping factors that are perceived to have the wrong sign for the partial regression coefficient (see comment above).

Response: We eliminated the approach of examining parameter signs and now report model fit statistics for all models that were evaluated.

The 20-day curve should be eliminated because the model is extrapolated beyond the range of the data. Fig. 2.1 indicates water transit time in LGR-MCN rarely if ever reaches 20 days.

Response: Water transit times were near or exceeded twenty days for much of the migration season in 2001. As such, the predictions in the figure are bounded by the observed range in the data.

Chapter 3

Page51 (lines 24-26, 33-36) and tables 3.2 (page 63) and 3.4 (page 74) – BPA Comment: Hydrosystem survival and system survival:

Response: In describing both hydrosystem and system survivals, it's clearly indicated that they aren't actual survival rates, and can exceed one. We disagree that they aren't useful in analyzing management options. Hydrosystem survival contains every hypothesized effect on overall survival of any particular proposed hydrosystem action in one term, and can be quite useful in modeling and simply in comparing expected changes in population growth rate due to management efforts in the hydrosystem. We do agree, however, that they aren't really used in the report, and since they can cause confusion and controversy among some readers, we will remove description and estimation of both quantities from the report. Since estimates of pathway probabilities then will not be used at all in Chapter 3, we will move description and estimation of these to Chapter 4, where they are used (for wild smolts).

Page 51 (lines 30-42) – BPA Comment: Assumption of density-dependent mortality needs more support and should be included here.

Response: This assumption related to system survival, which has been deleted from the 10-yr report.

Pages 61-79 (Part A) – BPA Comment: Bootstrap confidence intervals are not superior to theoretical normal theory confidence intervals arising from mark-recapture data analyzed with the CJS model.

Response: If we were only computing estimates of reach survival rates and collection probabilities, there would not have been the need for bootstrap confidence intervals and we would have simply used the theoretical normal theory confidence intervals. However, these parameters which we obtain from the CJS model are only components of more complex parameters. The estimation of number of smolts in categories T_0 and C_1 in LGR-equivalents uses CJS estimates of parameters S_2 and S_3 to expand LGS and LMN detection data, respectively, to starting values at LGR, while category C_0 uses estimates of parameters S_1 in addition to S_2 and S_3 in the estimation of starting smolts numbers at LGR. The estimates of smolt numbers in each study category are effectively combinations of the CJS estimates of S_1, S_2, and S_3 with tallies fish in cells of the reduced m-matrix, which are then divided into the tally of returning adults to obtain the study-specific SARs of SAR(T_0), SAR(C_0), and SAR(C_1). The ratio of SAR(T_0)/SAR(C_0) is used to estimate *TIR*s, and *TIR*s are multiplied by the ratio of S_R/S_T to arrive at *D*. Each of these computed parameters are a more complex function than the starting reach survival components produced with the CJS. The purpose of using bootstrap methods was to produce confidence intervals for these more complex parameters of interest.

Pages 61-79 (Part B) – BPA Comment: Show confidence intervals on all performance measures:

> *i). Geometric means of observed SARs, TIRs, or D values over years of study.*
>
> Answer: In the tables with SAR for each study category, the arithmetic mean and standard deviation is shown (not geometric mean, see the histograms of SARs presented in the 2006 CSS Annual Report), while those of *TIR* and *D* are geometric means. In each table, we will add parametric 90% confidence intervals about the average shown for parameter. Since parameters *TIR*, S_R, and *D* are log normally distributed, we will show confidence interval based on the anti-log of the arithmetic mean and confidence intervals of natural log transformed *TIR*, S_R, and *D*.
>
> *ii). Annual estimates of system survival estimates.*
>
> Response: This parameter is no longer presented in report.
>
> *iii). Annual extrapolated estimates of inriver survival (SR) from LGR to BON (Table D-21 to D-28):*
>
> Response: We will show the estimated 90% confidence intervals for the years with extrapolated estimates of SR with the caveat that those 90% confidence intervals may be narrower than what would have occurred if no extrapolation had been required.

Pages 61-79 (Part C) – BPA Comment: Bootstrap confidence intervals do not easily yield confidence intervals or stand errors on performance measures that are functions of other

parameters. Rather than report measures without some accompanying measure of uncertainty, standard errors or confidence intervals should have been computed in some way.

Response: With regard to the first sentence of this comment, it appears the reviewer did not understand how the bootstrap process was implemented. Given a release of N fish, each iteration of the bootstrap process was a random draw of N fish with replacement that created a new population of N fish for which all parameters of interest were computed. This process was repeated 1,000 times creating a distribution of 1000 observations for each parameter of interest. This distribution was sorted in order of increasing value, and the parameter value in positions 50 for lower limit and 951 for upper limit were selected for the 90% confidence interval. This approach does readily yield confidence intervals (as well as bootstrap standard errors), so it is unclear why the reviewer thinks bootstrap approaches "do not easily yield confidence interval or standard errors on performance measure that are functions of other parameters."

We are unaware of reasons why the bootstrap cannot be used to estimate confidence intervals (CIs) of quantities that are functions of other estimated quantities. It is true that standard errors of geometric means are easily calculated. However, it's not straightforward to estimate CIs of the geometric mean for short time series. In the special case where number of data points (years, in this case) is 1, the CIs will be lognormally distributed around the geometric mean. With many years of data the CI of the geometric mean approaches a symmetric (*t*-) distribution. However, with the short time series in the present analyses (6-10 years), the confidence intervals of the geomean are neither lognormally nor symmetrically distributed. We have not yet tried to develop an analytical method to estimate CIs of the geomean for short time series. Simulations could be used to estimate CI of the mean, however.

With regard to the second sentence of this comment, we do present the standard errors for the arithmetic means in the tables of annual SARs by study category. It was only in the tables with *TIR*, S_R, and D that we showed only the geometric mean. In our revision, we will show the 90% confidence intervals around the arithmetic mean or geometric mean as is appropriate for the specific table. See our responses to the above Part B portion of this BPA comment for additional details regarding this revision.

Pages 61-79 (Part D) – BPA Comment: Significant differences in point estimates are incorrectly based on non-overlapping 90% bootstrap confidence intervals. The reviewer states that significant differences may still occur even when two estimates have overlapping confidence intervals due to correlation between the two parameters as well as heterogeneity of variances between estimates of the two parameters. The review states that rather than look at the difference between SAR(T_0) and SAR(C_0), we should focus on their ratio TIR as the appropriate measure. The reviewer goes on to state that the determination of significant differences should be recalculated based on formal statistical test, and not on whether confidence intervals overlap.

Response: The review brings up valid points regarding correlation and heterogeneity of variance between the two parameter estimates, and states the *TIR* is the appropriate measure. In the report, we did not confine our investigation of significance to only differences between SAR(T_0) and SAR(C_0), but also indicated that when the lower limit of the *TIR* was greater than 1 there was evidence to statistically demonstrated significance higher SAR(T_0) than SAR(C_0). Based on the reviewers comments, we will revise the text to use the criteria of lower limit of non-parametric 90% confidence interval exceeds 1, which is effectively a statistical one-tailed

(α=0.05) test of H_0 *TIR* $\leq$1 versus H_A *TIR*>1 as the primary measure of whether SAR(T_0) is statistically greater than SAR(C_0).

Page 54 (lines 34-41 and 58-59) – BPA Comment: When inriver reach survival is not directly estimated to BON, you should use the term "extrapolated" instead of "expanded" since you are truly extrapolating past the available data. Did you looked at "per dam" extrapolation in addition to "per mile" extrapolation" You need to add standard errors or confidence intervals to the estimates of extrapolated S_R.

Response: The text will be revised to use the term "extrapolated" instead of "expanded" as recommended. We did compute extrapolations based on "per dam" as well as "per mile," but settled on "per mile" as the more appropriate method. In the reaches between LMN tailrace and MCN tailrace (76.6 miles) and JDA tailrace and BON tailrace (65.86 miles) there are two dams, so the two approaches produce similar results. However, in the MCN tailrace to JDA tailrace (73.94 miles) there is only one dam and a distance similar to the other two reaches noted above. Given this disparity between distances and number of dams involved, we believe the "per mile" extrapolation is more appropriate. We have added confidence intervals to the estimates of extrapolated S_R in Appendix D (as previously stated in responses to BPA Comments on pages 61-79 parts B and C).

Page 58 and 63 (lines 16-18) – BPA Comment: CSS includes steelhead jacks in SAR computation due to steelhead jacks having a fairly stable rate of return, while not including Chinook jacks in SAR computations due to Chinook jacks having a variable return rate. Removing jacks from the analysis because of their questionable contribution to spawning is understandable, but not because of a "highly variable jack return rate.

Response: The CSS report does state that the highly variable Chinook jack return rate among the various hatcheries versus low rate among wild Chinook was one reason for not including jacks in the SAR computations. The other reason, not stated though, is that jack Chinook are considered as having very limited contribution to spawning. We agree with the reviewer that our original sentence about the variable Chinook jack return rate seems out of place, and have deleted it from the text. However, we did not make any statement about steelhead jacks having a fairly stable rate of return. Instead, we simply stated in the methods section that we used 1-, 2-, and 3-ocean returns of steelhead. We will modify the methods section to say "Chinook jacks are excluded due to limited contribution to spawning."

Page 58 (lines 16-18) – BPA Comment: Conclusions (about D averaging 0.5 for hatchery and wild Chinook in recent years) are being presented pre-maturely and inaccurately in the methods section; and that these statements belong in the discussion section with corrections and justification. The reviewer points out that only 3 of 36 point estimates of D were $\leq$ 50% for hatchery Chinook in tables D-22 through D-26.

Response: We agree with the reviewer that the sentence about D averaging 0.5 does not belong in the methods section. Also, the statement that D was averaging 0.5 applied to wild Chinook only. The 10-yr geometric mean (excluding 2001) was 0.49 for wild Chinook with point

estimates in 6 years below 50% and 5 year (including 2001) above 50%. The text will be corrected to reflect this change, and moved to the results section.

Page 58 (lines 26-34) – BPA Comment: Measures S_R and S_T are called "hydrosystem survival," but these are not the hydrosystem survival described on pages 51, 59, and 60. Review wants our intentions explained or clarified.

Response: The "hydrosystem survival" as described on pages 51, 59, and 60 has been deleted from the report. With regard to the measures S_R and S_T, the text will be modified to state: "Therefore, to estimate $SAR_{BON\text{-}to\text{-}LGR}$ from $SAR_{LGR\text{-}to\text{-}LGR}$ for inriver migrating and transported fish, the effect of mortality through the hydrosytem must be removed by factoring out the survival rate from LGR to BON (S_R) for inriver migrants and survival rate in the barge adjusted for the inriver mortality incurred in order to reach transportation sites below LGR for the transported fish (S_T, see Formula 3.10 below)."

Page 59 (lines 13-21) – BPA Comment: Measures SAR_{T1}, SAR_{T2}, SAR_{T3}, SAR_{C1}, SAR_{C2}, and SAR_{C3} need formal definitions, both verbal and mathematical. Also, new notation C1, C2, C3, T1, T2, and T3, is used and needs definitions.

Response: Since we will drop the presentation of hydrosystem and system survival, these quantities will not be used in Chapter 3 and will be deleted. In Chapter 4, where these quantities will still be used, we will clarify their notation and description.

Pages 61-78 and Appendix D (Part A) – BPA Comment: Neither the actual numbers of tagged smolts transported from each dam nor the sample sizes used in the analyses are reported. The review states that this information is necessary for a complete and accurate peer review.

Response: The number of PIT-tags released in each year by species and rear type are presented in Appendix D. Tables D-1, D-3, and D-4 have a column labeled "total PIT-tags" which shows the total tag release each year and analyzed in the CSS for wild Chinook, wild steelhead, and hatchery steelhead, respectively. The actual number of PIT-tagged fish transported are included in Tables D-45 through D-47.

Pages 61-78 and Appendix D (Part B) – BPA Comment: It is unreasonable to assess the effectiveness of transportation based on small transport groups, even if they are augmented by the LGR equivalent approach.

Response: It must be noted that expanding the number of PIT-tagged fish released from LGS and LMN by the in-river survival rates between LGR and those downstream sites is not done for the purpose of augmenting the total transport number. It is necessary when indexing both transported fish and in-river migrants from LGR to expand the downstream counts to account for the fact that some fish die in route to the downstream transportation sites. As shown in Ryding (2006, see Appendix C), there is the need to properly apportion the mortality occurring between LGR and the downstream transportation sites to the transport and inriver study groups in order to obtain unbiased estimates of *TIR*s. We realize that small transport groups limits the ability to show significant differences between transported fish and in-river migrants in many years when

the goals of researchers was to return all PIT-tagged fish to the river at the transportation sites. But comparing trends between transported fish and inriver migrants over the years is providing evidence of the level of effectiveness of transportation as a mitigation tool for increasing SARs for wild Chinook and steelhead. The Chapter 4 methods explicitly deal with the effects of small sample size. They produce an estimated mean weighting by sample size, thereby accounting for small sample size.

Pages 61-78 and Appendix D (Part C) – BPA Comment: You should show project-specific TIRs; they are used in estimation of annual SAR in the body of the report, but are not specified.

Response: We assume that the second part of this comment applies to the annual estimates of overall SAR reported in this report. The overall SAR is computed by taking the study-specific SAR of groups T_0, C_0, and C_1 and weighting these SARs by the estimated proportion of fish in the total population (untagged and tagged) represented by each study-specific SAR. There is no use of *TIR* in this estimation. The results of evaluations of project-specific *TIR*s are covered in Chapter 4.

Page 61 (lines 26-28) and page 68 (lines 14-15) – BPA Comment: It is unreasonable to say that 2004 SAR is "low" at this point, since the 2004 returns are incomplete.

Response: With 3-ocean returns accounting, on average, for about 30% of the total adult return, and the SAR for the 2004 wild Chinook (based on the 2-ocean return) estimated at 0.30%, 0.31%, and 0.18% for categories T_0, C_0, and C_1, it was obvious that even when the 3-salt returns are added, the resulting complete return will provide "low" SARs.

Page 61 (lines 30-32) – BPA Comment: A reference made in the ISAB review of the 2005 CSS Annual Report refers to the NOAA finding that PIT-tagged survival is less than untagged survival. If the NOAA finding is true, then comparing SARs from PIT-tagged fish to target values is unreasonable unless we know the size of the bias introduced by tagging or tag loss.

Response: We address this issue in detail in Chapter 5. This line of reasoning assumes that the run reconstruction approach is correct. However, it may be that the difficulties in applying that approach has created SAR estimates that are too high. The "true" population SAR may lie somewhere between the levels estimated by these two methodologies.

Page 62 (line 13) and Table D-21 – BPA Comments regard the use of geometric mean to summarize point estimates of SAR, TIR, and D across years.

> *i). Use of the geometric mean needs justification, especially considering past criticism and the fact that the geometric mean will always be lower than the arithmetic mean.*
>
> Response: This same comment was made by BPA on the 2006 CSS Annual Report. The response was given on pages 170-171 of that document. In general, SARs for each study category approximate normality, as do the individual reach survival rates computed by the CJS method. However, the parameters S_R (i.e., the product of $S_2 \cdot S_3 \cdot S_4 \cdot S_5 \cdot S_6$), *TIR*, and *D* each appear to be lognormal distributed with skewness to the right. For these reasons,

the arithmetic mean was used for parameter SAR and the geometric mean was used for the other log-normally distributed parameters.

The geometric mean is a better measure than arithmetic mean of central tendency for right skewed (log-normally distributed) distributions such as *TIR* and *D*. They both represent ratios of survival rates, for which the ordering (i.e. which is numerator and which denominator) is arbitrary. From Zar (1984, p. 24): "[The geometric mean] finds use in averaging ratios where it is desired to give each ratio equal weight".

ii). Standard errors or confidence intervals need to be reported for the geometric mean (see earlier comment and suggestion.

Response: We have added standard errors and confidence intervals to the geometric means presented in Appendix D. However, the methods used for calculating these confidence intervals with short time series may not be appropriate, as discussed in response above.

iii). Low precision on D and TIR casts doubt on conclusions based on the geometric mean, especially those based only on a point estimate.

Response: We agree that low precision on annual estimates of *D* and *TIR* suggests that an unweighted mean should be interpreted cautiously. However, in the presence of large differences in mean values from target values, some inferences may be in order. The variable precision among annual estimates was a prime motivating factor in applying the methods used in Chapter 4, which allow stronger conclusions about the central tendencies of these quantities. As the number of years increase, the precision of geometric means will improve.

The reference to some estimates being only point estimates appears to refer to the parameter S_R and not *D* and *TIR*. As stated in earlier responses, we will show the 90% confidence interval for those S_R values that were extrapolated from a shorter reach, with the caveat that these confidence intervals will generally be narrower than would have occurred if sufficient data had been available to directly estimate the reach survivals in those lower reaches affected.

iv). The geometric mean inherently dampens the effect of extreme values, so the policy of excluding 2001 values from the geometric mean needs further justification.

Response: Excluding 2001 from the geometric mean was not a policy action. The drought conditions of 2001 were so unlike the other years that it was of interest compare the resulting *TIR* and *D* estimates of 2001 to the geometric mean of the other years. Data from 2001 were included in all estimates of *TIR* and *D* distributions made in Chapter 4

Page 66, 70, and D-17 (Tables D-29 and D-30) – BPA Comments: Annual SAR.

i). Annual SAR is discussed often and is described in word, but is never defined formally. An equation is needed to see exactly how the various components are incorporated.

Annual SAR values should be reported in a table for all species and stocks, with confidence intervals or standard errors.

Response: Coverage of the SARs described on pages 66 and 70 has been moved to Chapter 4, where equations and tables of results are presented.

ii). It would be useful to compare the annual SAR values to a simple ratio of the number of adults at LGR divided by the number of juveniles at LGR

Response: We disagree because the study fish do not migrate through the hydrosystem via the different routes in the same proportion as the untagged fish. Therefore weighted SARs are necessary.

iii). Tables D-29 and D-30 – BPA Comments that these tables should be explained clearly in text, using precise equations and clear definitions. It is unclear how the values reported here were defined, estimated, and used to compute the annual SARs. It is unclear what the S's mean, and what reaches they apply to. It is unclear where the covariances come from. No comparable tables were provided for hatchery fish.

Response: Appendix D presents information relevant to the whole document, not just to Chapter 3. These tables refer to work presented in Chapter 4, not Chapter 3. Apparently the commenter assumed they described an analysis in Chapter 3. Table D-29 is referenced in Chapter 4, and nowhere else. A reference to Table D-30 was inadvertently omitted from Chapter 4 and has been added. The purpose of the tables is clearly labeled in their captions; namely, to estimate covariance between pathways to estimate overall SAR mean and variance. The exercise was performed only for wild fish. The *S*s are also clearly defined in the captions. Moving the pathway probability language from Chapter 3 to Chapter 4, where the tables are referenced, should make the purpose of the tables obvious.

Page 67, Figure 3.7; page 70, last paragraph – BPA Comment: Figure 3.7 shows that the trend in SAR for wild fish over 2- or 3-yr time periods mimics the trend in SAR for certain hatchery stocks. However, Figure 3.7 also shows that SAR for wild fish did not closely track SARs for any single hatchery throughout the entire time period considered. It is therefore uncertain which single hatchery could be used as a surrogate for wild fish in future years. Also no error bars are provided on Figure 3.7.

Response: We agree that no one hatchery mimics the trend in overall SAR of wild Chinook, nor for trends in S_R (Figure 3.8), *lnTIR* (Figure 3.9), and *ln*D (Figure 3.10). That is why we do not make any recommendations for using only one hatchery as a surrogate. As for the lack of error bars in Figure 3.7, we present the 90% confidence intervals in Appendix E for the overall SAR parameter as "tot_sar." This appendix was not available at the time of the review. With up to 5 to 6 curves shown in Figure 3.7 across the years, the inclusion of error bound on each would have been too cluttered.

Pages 67 to 78 (Figures 3.7, 3.8, 3.9, 3.10, 3.13, 3.14, 3.15, 3.18, 3.19, and 3.20) – BPA Comment: Confidence intervals or standard errors are needed on these graphs.

Response: The goal of the figures was only to show the trends across years for the groups of fish being compared, and not to test whether significant differences occurred. We present 90% confidence interval in Appendix E for the overall *SARs*, S_R, untransformed *TIR*, and untransformed D. The 90% confidence intervals may also be found in Appendix D tables D-21 through D-28 for S_R, *TIR*, and *D*. Since the goal of the plotted data in these figures was aimed at only comparing trends over years, the error bounds about each curve was omitted in order to keep the plot uncluttered.

Chapter 4

Pg 81, 3rd paragraph – As the SARs are calculated in Chapter 3, they certainly do not have a binomial sampling variance, for both the numerator and denominator (i.e., C0 fish) are estimated random variables. For a binomial variance to be true, the denominator of the SAR would have [to] be known without error

Response: The numerator for SARs of any group is number of adult returnees detected at LGR, which is a count and not a random variable. It's true that the denominator for SAR of C_0 is an estimated quantity; however, as indicated later in Chapter 4, the CVs are small, and as demonstrated below, the deviation in variance from a true binomial is minimal. Similar methods of removing binomial variance from survival rate estimates which are not strictly binomial processes have been used. For example, Morris and Doak (2002) present an example using Kendall's (1998) beta-binomial method with data from desert tortoises: "[T]he capture-recapture method used to estimate survival doesn't yield a directly observed sample size. Instead, we used a rounded estimate of the total number of individuals that would have produced the observed number of live tortoises seen at the end of each time period, given the estimated survival rate" (pgs. 266 and 270.)

The variance of the ratio of returning adults to estimated number of smolts can be derived using the delta method, assuming both the numerator and denominator are random variables. A close approximation of the variance of the ratio of two random variables X and Y is (after Blumenfeld 2001, Eq 2.29)

$$Var\left(\frac{X}{Y}\right) \cong \left(\frac{\mu_X^2}{\mu_Y^4}\right)\sigma_Y^2 + \frac{\sigma_X^2}{\mu_Y^2} - 2\left(\frac{\mu_X}{\mu_Y^3}\right)\rho\sigma_X\sigma_Y,$$

where μ and σ^2 are mean and variance, respectively, and ρ is the correlation between X and Y. In the true binomial, variance of Y is zero, and the variance of the ratio reduces to the usual formula for variance of a binomial proportion *p*, i.e. $p(1-p) / N$, where *N* is the number of trials (number of smolts). By plugging in a value for CV of *N* when *N* is not known with certainty, the expected true sampling variance can be estimated. As noted in the discussion of Chapter 4, CVs of the estimate of C_0 are generally 2-4%. Below, we explore the effect of a CV of 4% in the numerator, along with two assumptions about the correlation between smolt numbers and adult returns (ρ), and two assumptions about mean smolt numbers, which reflect most of the range in

annual C_0 estimates. Mean SAR is assumed to be 1%, which is close to estimated values of SAR(C_0) for both wild steelhead and wild Chinook.

Table 1 shows that the effect of observed levels of variance in the denominator of SAR(C_0) is minimal. Simulations of binomial draws from a normal random variable representing C_0 indicate that, as expected, correlation between adult returns and smolts numbers increases with smolt numbers. Even at 5000 smolts, however, the estimated correlation at CV of C_0 = 4% is only 0.27, suggesting that the actual sampling variance departs little from the assumed binomial variance. Additionally, a positive correlation between smolt number and adult returns results in the binomial variance *over*estimating the sampling variance. This suggests that assuming binomial sampling variance may result in slight *under*estimation of environmental variance, for the range of correlations pertaining in this analysis. An expanded version of this analysis has been added to Chapter 4.

Table 1. Effect of CV of 4% in C_0 estimate on sampling variance of SAR(C_0), for different correlations and mean smolt number. SAR assumed = 1%. Binomial variance was assumed in Chapter 4 analyses. CV of SAR is sqrt (variance) / 1%.

Mean C_0	ρ	Actual variance	Actual CV	Binomial variance	Binomial CV
200	0	4.97×10^{-5}	70%	4.95×10^{-5}	70%
200	0.5	4.68×10^{-5}	68%	4.95×10^{-5}	70%
5000	0	2.14×10^{-6}	15%	1.98×10^{-6}	14%
5000	0.5	1.58×10^{-6}	13%	1.98×10^{-6}	14%

Page 82, lines 15-17. Akcakaya (2002) is cited as a foundation for the method used to remove sampling variance to estimate environmental variance. The method presented in Akcakaya (2002) is appropriate for census data, but not for mark-recapture data, such as the data analyzed in this report. Akcakaya (2002) refers to both Burnham et al. (1987) and Gould and Nichols (1998) for variance-components method of removing sampling variance from mark-recapture data (see below, comment on pages 82-87).

Response: Gould and Nichols (1998) point out that if the population parameter is known (directly observed), there is no variance component associated with sampling error. Gould and Nichols' analysis considered cases with two sources of "sampling" variability (pg. 2532): 1) variation associated with the inability to count at sampling period $i + 1$ every marked survivor from period i, and 2) demographic stochasticity producing binomial variation in the number of marked survivors at the end of period $i + 1$. In the present analysis, there is no sampling variance of the first kind. All (or nearly all) surviving adults are "captured" by PIT-tag detection at LGR, i.e. there is a "census" of survivors. Therefore, since the present analysis deals only with demographic stochasticity, the more involved methods of Gould and Nichols for estimating the first kind of sampling variance and its covariance with the second kind are not required.

Page 81, 4th paragraph – The belief that there is a single probability distribution of SAR, TIR, or D over a long time period assumes that there is no temporal trend in the measures, such as may

be caused by global climate change. Chapter 3's focus on trends in these measures suggests an assumption that the measures are changing over time, which is inconsistent with the assumption that they arise from a single beta distribution

Response: CSS's primary purpose is in data collection and monitoring, and in particular estimating SARs and the efficacy of smolt transportation. Using the presents methods to estimate distributions reflecting inter-annual variation in SARs and their ratios observed in the recent past requires no beliefs about the factors influencing SARs. In the introduction of Chapter 4, the assumption under which the distributions derived would be useful for prospective modeling is clearly stated (pg. 81). In any system, the future cannot be guaranteed to be identical to the past, yet there is no end of literature presenting estimates of recent population abundance, survival rates, population growth rates, etc, in an attempt to understand the current state of the system. Describing *what* has occurred is not inherently inconsistent with exploring hypotheses about *why* it occurred.

Page 86, lines 19-22 - Equation(4.4) for the variance of a product applies only for independent random variables. This equation cannot be used to calculate the variance of a product of inriver survivals over adjacent reaches (i.e. S_R,), because these survival estimates are correlated as based on the CJS model. Instead, the delta method (Seber 1982:7-9) should be used.

Response: The paragraph immediately under Equations 4.4 and 4.5 indicates that the assumption of independence of the random variables is required. Here and elsewhere this assumption is made, evidence supporting its reasonableness is provided. Description of the accuracy of the bootstrap procedure in reproducing overall variance in S_R from individual reach survival rate estimates which covary is presented elsewhere in these responses.

Page 82-87 - Kendall's (1998) method is a parametric approach to variance component estimation that makes unnecessarily restrictive assumptions, i.e., a. Measurement error is binomially distributed. b. SARs are beta-distributed (and following equations and numbered points).

Response: The commenter has misunderstood the method of variance partitioning used. As explicitly stated in the Chapter twice (pg. 82-83), Akçakaya's (2002) method of variance partitioning, rather than Kendall's, is used. On page 82 we explicitly note that Akçakaya's method is an alternative to the approach of Kendall. Akçakaya's method involves no assumption about the form of distribution of the resulting survival probability. Our rationale for representing the resulting environmental variances with beta distributions is provided elsewhere in Chapter 4. The goodness of the assumption of binomial sampling error of C_0 SARs is discussed above.

There are several implications of the parametric approach taken to variance component estimation used in the CSS report, including the following: (following bullet points).

Response: See Chapter 4 for rationale for choosing beta distributions to represent variability in SARs. Kendall (1998) and Morris and Doak (2002) use similar methods to estimate beta

distributions to describe variability in survival rates. See these authors for more detailed rationale, and survey the literature on stochastic population modeling for numerous examples of using the beta to represent variation in survival probability. Those authors do not expect proof that the limited data in hand in most conservation problems conforms to a beta distribution (this is impossible with the short time period data sets available in most conservation problems—it would take many, many years of data to allow discrimination between beta and alternative distributions). Perhaps the commenter could suggest a different probability distribution that would better reflect variation in survival rates over many years.

The exact form of the beta distribution used is presented in Chapter 4—it is the identical form as used by Kendall (1998), as referenced. Equations 4.2 and 4.3 (4.3 and 4.4 in the revised Chapter) show how the parameters of the beta distribution are derived from the mean and environmental variance derived from using Akçakaya's approach. The commenter's equations 5 and 6 can be derived from equations 4.2 and 4.3 by solving for mean and variance; or the converse operation can be performed.

Where data are sufficient for plausible estimates of correlation, our analyses do not assume that SARs of different groups vary independently. In fact, in estimating *TIR* and *D* distributions for Chinook (where data are sufficient for estimation of correlation), we include covariance between transport and in-river groups. With regard to global warming, see earlier response to page 81 comment.

Page 88, Table 4.1 • The estimated demographic variance is greater than total variance, suggesting something is wrong and thus casting doubt on all methods and results in this chapter. • Observed correlations between point estimates of SAR for transport and C0 groups for wild steelhead are explained by small transport groups and so are not used. However, such small transport groups (we are not told the actual sizes) produce unreliable parameter estimates that can seriously distort interpretation of results.

<u>Response</u>: Gould and Nichols (1998), which the commenter commends, produced negative estimates of variance (due to estimated sampling variance being > total variance) for a number of their sample data sets. They reference literature indicating that negative estimates of variance are not uncommon in the variance components literature (pg. 2534-2535). In the CSS study, the one case of estimated sampling variance slightly exceeding total observed measurement (steelhead transported from LMN) is a consequence of large sampling variation due to only 8 PIT-tagged adults returning to LGR over the 6 years. In this case, a reasonable and conservative approach is to use the observed inter-annual variance as an estimate of environmentally driven variance.

Uncertainty in parameter estimates is explicitly estimated and accounted for in these procedures. The effects of small sample sizes combined with low SARs can be seen in the resultant wide confidence intervals for SARs of LGS- and LMN-transported steelhead (Figure 4.2). The effect of this uncertainty is carried into estimates of *TIR* and *D* for these projects and explicitly presented. Assuming independence of SARs in estimating *TIR* and *D* distributions is a reasonable and conservative default procedure in this case. The relevant raw data, including numbers of PIT-tagged fish transported from each project, and detected upon return as adults at

LGR, can be found in Appendix E of the 2006 CSS Annual Report. The raw data were also provided in electronic form to BPA at the start of the comment period.

Chapter 5

BPA General Comments, p. 2: The CSS continues in Chapter 5 its comparison of upstream and downstream Chinook salmon stocks. As in the past, multiple upstream hatcheries and collection points are used, while only a single downstream hatchery and collection point (for wild fish) is used, despite the ISAB's recommendation to incorporate more downstream stocks. Given that this is a retrospective report, it is understandable that the CSS could not immediately include additional downstream stocks. While the CSS does perform useful comparisons of biological characteristics of the upstream and downstream stocks, their upstream-downstream analysis is invalid in other critical ways. The CSS uses an invalid performance measure to identify delayed mortality caused by the hydrosystem. This approach assumes no natural mortality for smolt should occur between upstream and downstream sites. When the performance measure is corrected for the longer migration of the upstream stocks, there is little or no evidence of delayed hydrosystem mortality for hatchery Chinook salmon. Similarly, the CSS Report does not consider the longer distance to travel for upstream stocks when comparing travel and arrival times of upstream and downstream stocks. Even if the hydrosystem were not in place, the upstream stocks would still have farther to travel than downstream stocks.

Response: To clarify, three downriver populations are included as an aggregate in the analysis for wild Chinook: North Fork, Middle Fork and upper mainstem John Day Rivers. We have noted both the ISAB recommendations and the CSS proposals to increase the number of downstream wild and hatchery populations, which BPA has yet to fund.

The BPA reviewers appear to be confused on the purpose of the upriver/downriver analysis, which was stated in the Chapter 5 introduction (p. 106): "*our specific interest … is whether upriver/downriver differences in SARs for wild and hatchery stream-type Chinook were consistent with the differential mortality estimated from SR [spawner-recruit] models for wild populations"* Previous published SR analyses indicated there was a systematic increase in mortality for Snake River populations, which did not occur in the downriver populations, associated with the construction and operation of the FCRPS (e.g., Deriso et al. 2001; Schaller et al. 1999). In the SR model formulations, any differences in smolt mortality caused by different travel distances would be incorporated into the intrinsic productivity (Ricker "a"). Obviously, the migration distance for upriver and downriver populations did not change over the time period of FCRPS development; Water Travel Time (WTT), Fish Travel Time and hydro impacts did change with this development. WTT for Snake River stocks before FCRPS development were only about 2-3 days; Snake River smolts were historically able to arrive at the estuary more in synchrony with their morphological, physiological and behavioral development (e.g., Budy et al. 2002; ISG 1999). Available evidence from a mostly free-flowing migration corridor (Whitebird trap on the Salmon River to Ice Harbor Dam) also suggests smolt survival was high before FCRPS development (Raymond 1979). Applying the survival per mile from the Raymond study the information suggests that the historic survival from Lewiston to Bonneville dam was over 90%.

The BPA reviewers appear to be confusing differential mortality and delayed mortality. The analysis in the CSS report estimated differential mortality based on SARs to compare with differential mortality estimated by SR analyses (see equations 5.2 and 5.3). We did not explicitly estimate *delayed* mortality for in-river migrants, although the upriver/downriver SAR differential mortality comparisons are relevant to such an analysis.

• *On a yearly basis, p should be estimated as (i) Manly-Parr formula, (ii) CJS formula, but not the CSS formula p = (N detected at BOA) / (N detected at BOA + N passing BOA undetected that were later detected upriver).*

Response: From the reviewer's comment, it was apparent that the formula shown in footnote 5 of Table 5.9 caused a misunderstanding of our approach. That footnote has been corrected. Our approach is identical to what the CJS model produces in a three site model – site 1 for release, site 2 for BOA and site 3 for pool of upriver dams. In the Burnham et al (1987) monograph, the estimate of collection efficiency at site 2 is $p_2 = m_2/ m_2 + z_2 \cdot (R_2/r_2)$. With only 2 recovery sites, this equation simplifies to the following using the reduced M-matrix in the Burnham monograph for k=3 sites:

Cohort	Site 1 Release	Site 2 BOA	Site 3 Upstream	Sum detections
1	R_1	m_{12}	m_{13}	r_1
2	R_2		m_{23}	r_2 .
	Column sum	m2	m3	

Sum for z_2 is m_{13}
Sum for r_2 is m_{23}
$R_2 = m_{12}$ since there are no removals at BOA
$m_2 = m_{12}$

Formula for $p_2 = m_2/ m_2 + z_2 \cdot (R_2/r_2)$
$= m_{12} / (m_{12} + m_{13} \cdot (m_{12} / m_{23}))$
$= m_{23} / (m_{23} + m_{13})$

The number of fish in m_{23} = *N* jointly detected BOA & upriver
and number of fish in m_{13} = *M* passing BOA undetected & detected upriver

Substituting these equalities gives the formula that we are now showing in footnote 5 of Table 5.9. Therefore, we are actually utilizing the CJS model approach and producing a valid estimate of p_2 at BOA.

BPA comment: Page 106, lines 11-22
• *Critiques of the single release-recapture (SR) analysis and PATH have demonstrated the reliance of latent mortality results on untestable assumptions, e.g., stock-specific Ricker a's*

versus a common Ricker a. Additionally, climate effects have been shown to account for the majority of latent mortality. These criticisms should be addressed in this chapter.

Response: The BPA reviewers seem to be confusing delayed or latent mortality with differential mortality; also SR is the abbreviation for spawner-recruit, not single release-recapture (see p. 115). The differential mortality estimated from PIT-tag SARs (equation 5.3) can be used ultimately to test differential mortality estimates using different SR (spawner recruit) model formulations. It is important to note that the reviewers are criticizing the published material we referenced, however, we did not perform SR analysis in the CSS report. The purpose of the CSS PIT-tag analysis was to provide independent estimates of differential mortality, for comparison with estimates from published SR analyses (Schaller et al.1999, Deriso et al. 2001, Schaller and Petrosky 2007). We are aware of one alternative SR model that suggests differential mortality may be low, which uses a common Ricker "a" for all populations (R. Hinrichsen, unpublished manuscript); other models investigated by Hinrichsen yielded differential mortality estimates similar to that in Figure 5.16 in the CSS report. Given the 4-fold difference in SARs estimated between Snake River and downriver populations, the common Ricker "a" hypothesis does not appear very plausible. Other issues with this hypothesis include the habitat quality differences among Columbia Basin streams (and thus expected differences in intrinsic productivity) and the fact that the common Ricker "a" formulation produces other questionable parameter estimates. Regardless, by continuing and expanding CSS PIT-tagging of upriver and downriver populations more formal testing will be possible through analyzing these SAR estimates.

BPA comment: Page 106, lines 19-20
• It is not explained and it is unclear how direct mortality, differential delayed mortality of transported smolts, and the common year effect were accounted for in the SR comparisons.

Response: We provided three references which provide detail regarding how delayed mortality of in-river migrants may be partitioned from total mortality. Since we did not explicitly estimate delayed mortality in this report, we did not provide equations from these literature sources that did make delayed mortality estimates.

BPA comment: Page 107, line 26
• "Overall SAR" is never defined, either here or elsewhere in the report. Presumably it is equal to "annual SAR," which is also never defined analytically.

Response: We added the definition of overall SARs and a reference to the detailed analytical description in Appendix B (see page B-10 – *Estimation of overall annual SARs*).

BPA comment: Pages 110-111
• Run Reconstruction SARs: Include jacks and adults; measure returns to mouth of Columbia River.
• CSS SARs: Include only adults (Chinook), no jacks; Measure returns to LGR
• Are run-reconstruction SARs and CSS SARs really comparable? It has not been justified that direct comparison of the measures is appropriate.

Response: We modified the language to indicate that both run reconstruction and CSS SARs in this analysis represented returns to the uppermost dam (Lower Granite since 1975) adjusted to account for harvest. Our initial comparison had the (quantitatively minor) inconsistency that we included jacks in the run-reconstruction estimates, which we have fixed.

BPA comment: Page 112, lines 15-19: How is WTT defined?

Response: We added the following language: Water velocity in the mainstem migration corridor is generally expressed as the average time (in days) it takes a water particle to travel through a river reach (water travel time) during a specified period.

BPA comment: Page 114, lines 3, 9; Figures 5.5, 5.6
• What does "frequently incorporated in multiple regression models" mean?

Response: We changed "incorporated" to "selected" in the caption.

BPA comment: Page 115, Multiple Factor Model, lines 5-31
• How were candidates for independent environmental covariates selected? What were they? Only WTT, PDO, and an upwelling index are named, and it is unclear whether other covariates were considered.
• Harvest and temperature are known to affect SARs and do not appear to have been considered.
• Were any other "inriver" predictors than WTT considered?
• Were interaction terms considered in the multiple regression models?
• Typo in SAS version (presumably 9.1, not 91).

Response: Candidates for independent environmental covariates were those that have been previously linked to, or hypothesized to influence, salmon SARs (p. 112). Other potential juvenile migration variables covariates for future analyses may include a measure of spill proportion and proportion of the run transported. Because SARs in this analysis represented pre-harvest adult recruits, harvest was already accounted for. We did do some exploratory analysis with average monthly sea surface temperatures at various latitudes. However, it was not very informative and we believe sea surface temperature was incorporated by the PDO, a large-scale index of sea surface temperature anomalies in the North Pacific Ocean. We did not include any interaction terms, although, this may be attempted (for the longer time series) in future analyses. The SAS version typo was corrected.

BPA comment: Methods: Snake River and Downriver SAR Comparison (pp. 115-119)
• There has been much previous criticism of the upriver-downriver comparisons made by the CSS and of the spawner-recruit model used to justify the upriver-downriver comparisons. Insufficient response has been made to these criticisms.

Response: We went into detail addressing each off the past criticisms for the upriver-downriver approach on page 119-120. We focused on the published upriver-downriver criticisms and the published responses to these criticisms. In addition, we provide a summary of analyses comparing biological characteristics of the two population groups.

Page 116
• Lines 7-8
− How is μ_t defined and estimated? Provide an equation showing how value is calculated. Is this the same μ as in Eq. 5.3, or is it the differential mortality defined verbally based on Eq. 5.2?
− The "delta model" should be defined.

Response: We did not estimate μ_t from SR data in the CSS report, we only compare PIT-tag estimates of differential mortality to previously published estimates of μ_t. We specifically referred the reader to Deriso et al. (2001) equations 4-6 for estimation of μ_t. The delta model was defined as the primary model in Deriso et al. (2001) just above equation 5.2 (p. 115 line 44).

BPA comment: • Equation 5.3: If there is no delayed mortality from hydrosystem, then we expect $\exp(-\mu_{SAR,t}) = S_{J(LGR-JD)}$. This important point is omitted from the report.

Response: The subscript "J" in the reviewer's comments is not clear to us. However, see our response to the reviewer's table 1. If we understand the reviewer's point, partitioning in-river survival ($S_{(LGR-JD)}$) from the SARs is not analogous to estimating differential mortality from SR data. Also, this formulation, as we interpret the reviewer's point, does not account for the large proportion of fish which are transported.

BPA comment: Page 117
• Line 18: Only a single hatchery (Carson) is used for the downstream hatchery Chinook salmon.

Response: The CSS study has only received funds to maintain a long time series of PIT-tag SARs for only one downriver hatchery.

BPA comment: Page 119, Table 5.9
• This table is very difficult to understand. The caption does not agree with the notation used in the table. Values reported in the table are not sufficiently explained. It appears that the formula used to estimate BOA detection efficiency (p) is wrong.

Response: Agree that caption does not clearly state the purpose of the table, so it has been revised to read:
" Table 5.9. Estimated PIT-tag detection efficiency of combined adult detectors at Bonneville Dam based on combined unique detections of PIT-tagged adults at McNary, Ice Harbor, and Lower Granite dams."

Also, footnote 5 was misleading as currently stated and has been revised to read:
" Calculated as p = (N jointly detected BOA & upriver) / (N jointly detected BOA & upriver + M passing BOA undetected & detected upriver)

• What are the values reported in the row "GRA, MCA, IHA?"

Response: The sum of unique PIT-tagged adults (≥ 2-ocean) detected at either IHA (Ice Harbor Dam, where IHA and ICH are possible detection site names), MCA (McNary Dam, where MC1

or MC2 are possible detection site names), and GRA (Lower Granite Dam). Each returning PIT-tagged adult is counted only once from this pool of three recovery sites.

• MCN and IHA are not mentioned in table caption.

Response: Caption has been rewritten so that all three dams are included

• The estimate of p based on detections at BON and upstream is INVALID if it is based on detections from different years, unless upriver adult survival to GRA is constant across return years, and detection probabilities at MCN, IH, and GRA are constant across return years. This is not true, so estimates of BOA detection efficiency presented here are invalid.

Response: The annual detection efficiency probabilities at BOA were estimated at the level of the smolt migration year, so as to allow a single expansion factor at BOA for total adult return counts. The reviewer's concern that upriver adult survival and detection probabilities may change across years is not a problem since we are creating the BOA detection efficiency as a conditional probability, given the sum of unique (counted only once) PIT-tagged adults detected above BOA. Since these fish are detected above BOA, we know they were alive when passing BOA, and so a conditional probability calculated as p = (N jointly detected BOA & upriver) / (N jointly detected BOA & upriver + M passing BOA undetected & detected upriver) is a valid approach.

BPA comments: Methods: Comparison of biological characteristics of Snake River and downriver smolts (pp. 119-121)
Page 120
• In general for upstream/downstream comparisons, was goodness-of-fit considered or examination of residuals performed? Show results.
• With only 6 years of data, this is not a long time series, which limits the amount of useful information that can be gleaned from it.

Response: It is not clear what reviewers are suggesting with the first comment. Six years of data are what we have available, however, sample sizes (numbers of tagged smolts) are large enough within and across years to detect statistical differences where they exist.

BPA comments Page 121
• Lines 13-14: No migration distance is given for JDAR1 fish. Comparison of survival and travel time between upstream and downstream fish should incorporate migration distance for the two groups of fish.

Response: Reviewers' comment is not clear; we presented the migration distances in lines 13-14.

BPA comments • Lines 40-41: Basing analyses on (Number of BON detects/Number released at trap) assumes that all groups have the same conditional detection probability at BON. This is likely to change with arrival timing.

Response: It is unclear what the reviewers point is?

BPA comments: Results: Overall SARs (pp. 122-127)
Page 122, lines 32-34
• "Removing sampling variability" resulted in lower mean SAR. Does this always occur?
Page 126, lines 17-19
• The CSS has been using a geometric mean previously, but here does not identify the type of mean used for mean SAR.
• It is unclear what the reference to the t-distribution means. If a formal t-test is being performed, this should be stated simply. Note that while these arithmetic means may be compared using a t-distribution, the geometric mean should not.

Response: The variation portioning ("process error") method used in Chapter 4 uses a weighted mean SAR, which usually will differ from the unweighted mean. The amount and direction by which they differ depends on how sampling error is distributed among years with varying point estimates of SARs.

In the draft report we did not log transform the SARs. In the final draft we recalculated the mean SAR based on natural log transformation and the percent of the distribution above 2%, and modified the text accordingly.

We did formally use a t-test and specifically stated our methods on page 107 lines 33-39.

BPA comments: Results: Relationship between SAR and environmental covariates (pp. 128-131)
Page 128
• Lines 4-8: The data for the PIT-tag SARs and environmental factors are not presented in this report.
• In general, references to figures should be proofread. There are mistakes in figure references throughout the chapter, making it difficult to follow the narrative.

Response: The CSS PIT-tag SARs (LGR-LGR) are in Appendix E (data was sent to BPA reviewers on request). We also cited the source of the run-reconstruction SAR data set, and provided the websites for environmental data. We corrected the figure references in the final draft.

BPA comments: Line 11: What is meant by "bi-variate results?" Is this regression of a single response variable on a single predictor variable? A vector response variable on one or more predictor variables? A single response variable on two or more predictor variables?
• Table 5.4: Did the CSS consider correlation between PDO and UP45n? Both types of measures are used in the same regressions, apparently.
Page 129, Table 5.5
• It should be explained why SepPDO is used rather than JulyPDO as a covariate, when JulyPDO looks better than SepPDO for both the long and current time series. Page 129, Table 5.5

• It should be explained why SepPDO is used rather than JulyPDO as a covariate, when JulyPDO looks better than SepPDO for both the long and current time series.
Page 130
Page 130
• In general for regression with environmental variables:
− What was the set of candidate predictor variables? Was it only PDOs, UP45ns, and WTT?
− How model selection was performed needs to be specified?
• Lines 13-14 – The report says that WTT was "less significant for the shorter time series," but Table 5.7 indicates that WTT was not at all significant if the model includes upwelling index (Table 5.7, Current Time Series).
Page 131, lines 1-2, and Table 5.7, Current time series
• What model selection criterion were used to identify the "best" model?
• The "best" model shows no predictor variable significant at the 10% level when upwelling index is included.
• Without upwelling index (NovUP45n), SNWTT and MayPDO become significant. Was multicollinearity between these parameters and NovUP45n considered? And how?

Response: Our use of the term bi-variate results refers to regression of a single response variable on a single predictor variable. Our primary concern with correlated independent variables was to screen against highly correlated monthly variables within the PDO (such as between April and May or May and June) or the within the upwelling indices. However, the correlation between SepPDO and NovUP45n model selection (ocean variables selected for the best 3 parameter model for the long time series), was negligible (-0.02). JulyPDO was screened out from the regressions because it was highly correlated with MayPDO and SepPDO (0.72 and 0.66, respectively; Table 5.4). MayPDO and SepPDO were not as highly correlated (0.46) as some other possible combinations. The list of candidate variables (after screening for correlated variables) included SNWTT, AprUP45n, OctUP45n, NovUP45n, MayPDO, and SepPDO. The model selection process was described in methods (p. 115). Text was modified to include the one non-significant result for the current time series. The time period we call current is a short time series, so the result that MayPDO and SNWTT became significant without NovUP45n should not be surprising. The correlation between SeptPDO and SNWTT was also very small.

BPA Comment: Results: Snake River and Downriver SAR Comparisons (pp. 131-136)
• The CSS upstream-downstream comparison of SARs is based on the performance measureU/D =$S_{LGR-BON}/S_{JD-BON}$. If there is no differential post-JD mortality for upstream fish, then we expect U/D to equal S_{LGR-JD}, inriver smolt survival from LGR to JD.

…numerous comments continued through…

BPA comment: It is obvious from Table 1 that the value of U/D (and by extension, SARμ) alone does not indicate whether or not "differential mortality" has occurred.

Response: The BPA reviewers seem to misinterpret the purpose of the SAR comparisons, which is to evaluate if the same patterns evident in published SR (spawner-recruit) differential mortality were present in SARs. The purpose was stated in the Chapter 5 introduction (p. 106): "*our specific interest … is whether upriver/downriver differences in SARs for wild and hatchery*

stream-type Chinook were consistent with the differential mortality estimated from SR [spawner-recruit] models for wild populations" Previous SR analyses indicated there was a systematic increase in mortality for Snake River populations, which did not occur in the downriver populations, associated with the construction and operation of the FCRPS (e.g., Deriso et al. 2001; Schaller et al. 1999).

The reviewers' comments contain a purported comparison of survival from John Day to Bonneville Dam with the ratio of SARs from upriver and downriver stocks (Table 1), and assert that this comparison would be more appropriate than a SAR comparison that indexes smolts leaving the production areas (i.e., at the first dam). There are two problems with the reviewers' approach. Their proposed approach is inconsistent with the original SR definition of differential mortality (e.g, Deriso et al. 2001), where spawners were indexed at the spawning grounds and recruits were indexed at the Columbia River mouth (p. 116, lines 29-31). Second, the reviewers propose to account only for the passage mortality experienced by in-river migrants and not that of transported smolts (the migratory route the majority of fish experience). One could, in theory, fix the smolt indexing location at any number of locations (JDA or BON), but this would be a very different analysis, and not consistent with the SR based estimates of differential mortality. It is not clear what the reviewers' proposed adjustment only for in-river survival would accomplish, other than further confuse this issue.

<u>BPA Comment</u>: Page 132, Table 5.8
• How are the SARs for downriver wild Chinook salmon estimated? If simple return ratio, why not use same method for Snake River fish?

<u>Response</u>: The methods for John Day wild Chinook SARs are described on p. 116-117 and Table 5.7. As explained in Appendix B, Snake River annual SARs required weighting by study category (T_0, C_0, C_1) to reflect their true proportion in the run-at-large. Because John Day smolts were not experimentally separated into different study categories, there was no need to perform this weighting for these fish.

<u>BPA Comment</u>: Page 134
• Lines 2-5: The CSS claims that the SAR to BON is always higher for the downriver (hatchery) fish, but that is not true for 2003.
• Lines 13-16: The reason given for not providing a confidence interval on SAR for downriver fish in 2004 is because an average survival to BON from previous years is used. However, that survival is not known without error, so a measure of uncertainty should be reported on survival to BON for 2004, and that error could be propagated to produce a CI on SARs.
• In general, the CSS addresses uncertainties incorporated by using a single downstream hatchery stock when the upstream/downstream results show no effect of the hydrosystem (i.e., for hatchery Chinook salmon), but not when the upstream/downstream results do imply hydrosystem effects (i.e., for wild Chinook salmon). This sounds like an inconsistent approach.
Page 135, Table 5.10
• In some years, upriver SAR > downriver SAR for hatchery Chinook salmon, despite additional inriver migration for upriver fish. Presumably, this result is unexpected and should be

addressed. Such results may be due to large measurement error that obscures the relationship or the upstream/downstream pairing is a mismatch.

Response: The draft text in question (p. 134 lines 2-5) does not claim the downriver hatchery SAR was "always" higher; we added the word "generally" to avoid misinterpretation in the final version. In the future, a CI for the SAR of downriver fish in 2004 could be generated with this measure of uncertainty. The point of the reviewers' comments about the single hatchery stock is not clear. The downriver wild aggregate is comprised of three populations, and CSS has proposed adding more populations to reduce uncertainty from this factor. We are simply noting on p. 134 that in addition to use of a single downriver hatchery stock, that use of hatchery fish as surrogates of wild fish performance has additional potential confounding factors: hatchery practices, disease, rearing conditions and fitness.

BPA Comment: Results: Comparison of Biological Characteristics of Snake River and downriver smolts (pp. 136-143)
Page 139
• Lines 2-3 It says that there is a significant (P < 0.001) difference in density-adjusted mean fork lengths of 106 and 106 mm (for IMNTRP and JDAR1), and separately of 100 and 100 mm (for SALTRP and SNKTRP).
• Lines 6-7: The report is inconsistent when it says 74 mm vs. 121 mm in fork length is not significant, especially considering that they previously defined any differences >5 mm to be biologically significant.

Response: As noted on page 120, because the sample sizes were very large we had the ability to detect small fork length differences with a high degree of statistical significance. We changed the text to more accurately reflect the results of comparing fish sizes for John Day to Snake Basin populations.

BPA Comment: Page 141, lines 11-13
• "Smolts from upriver populations and downriver-origin smolts migrated at a similar rate, once their different migration distances were accounted for." What does this mean? Their migration "rate" (i.e., distance traveled per unit time) already accounts for differing migration distances.

Response: We changed the sentence to say: This comparison demonstrates that smolts from upriver populations and downriver-origin smolts migrated at a similar rate.

BPA Comment: Page 142
• Lines 15-17 – The observation that upriver smolts took longer to travel to BON than downriver smolts is not surprising since they leave at the same time and travel at the same rate, given that upriver smolts have farther to travel.

Response: As discussed above, distance did not change as a result of FCRPS development; water velocity (and WTT) did change. In the impounded river system, smolts are moving at approximately the rate of water velocity (e.g., Fig. 5.22); current average WTT is about 19 days. WTT for Snake River stocks before FCRPS development was only about 2-3 days; Snake River

smolts were historically able to arrive at the estuary more in synchrony with their morphological, physiological and behavioral development (e.g., Budy et al. 2002; ISG 1999). If the optimal estuary entry timing for an individual smolt is 12 days after passing LGR, on average, it will arrive a week later than optimal, given the current FCRPS configuration and management.

BPA Comment: Results: SARs by Bonneville Arrival Timing (pp. 144-146)
Page 143, lines 1-2
• The "pattern of delayed arrival" was not consistent across years, as is stated – See years 2000 and 2003.

Response: We added the word "generally" to final draft.

BPA Comment: Page 144
• Lines 11-14 – What reference point is used to determine that upstream smolts experience delayed migration?
• Lines 16-18 – What does "significantly experienced lower SARs" mean? Does this mean that the difference or ratio between the SAR for wild upstream Chinook and wild downstream Chinook was statistically significant? Biologically significant?

Response: The reference point is the large reduction in water velocity from historical conditions discussed above, and strong observed relation between FTT and WTT (see also Chapter 2).

The sentence on line 16-18 was reworded: "All groups of Snake River wild Chinook experienced significantly lower SARs (Bonneville to Bonneville) than John Day wild Chinook within the same arrival time period and for the season, based on non-overlapping 90% CI." This difference in SARs would be statistically significant, and considering that the point estimates differ by about 2-fold, also biologically significant.

BPA Comment: Page 145, Figure 5.23
• Binomial confidence intervals are shown, but error is not binomial for C0, C1, and T0. Recalculate appropriately.
• In some years, large numbers of upriver migrants are omitted from the analysis by restricting attention to 16 April – 31 May window.

Response: The number of smolts arriving at Bonneville is a known quantity; therefore applying binomial confidence intervals is appropriate.

The purpose of this comparison was to compare SARs from the same arrival timing, therefore, because there were so few John Day smolts during the late arrival period it was omitted from the analysis. Note that all the data are available in table 5.16.

BPA Comments: PIT-tag SARs versus SAR of run-at-large (p. 147)
• Lines 3-5: Are run-reconstruction SARs and CSS SARs mathematically comparable? Justify.
• Lines 12-19: Assumptions necessary for the run-reconstruction SARs are discussed, but not assumptions for the CSS SARs.

Response: The run reconstruction SARs in the draft report inadvertently included jacks. This has been corrected to exclude jacks for consistency with the CSS SARs, and text has been modified. Methods and statistical assumptions for the CSS SARs are covered in Appendix B (and elsewhere) in the report, and the issue of a potential negative bias for PIT-tag SARs was addressed in this section and the discussion.

BPA Comment: Discussion (pp. 148-151)
Page 148
• Lines 21-22: The limitations of small sample size cannot be avoided by using multi-year methods, as indicated here. Multi-year methods result in conclusions that are based on many uncertain estimates (due to small yearly sample sizes), instead of based on only a single uncertain estimate. This simply expands the problem of small sample size.

Response: The text referenced refers to the analyses presented in Chapter 4. The Chapter 4 methods explicitly deal with the effects of small sample size. They produce an estimated mean, weighting by sample size, and so account for small sample size, rather than "expanding the problem". Sampling variance is estimated and removed from total variance to get a truer estimate of actual inter-annual variance in SARs, and hence in the ratio of SARs as well.

BPA Comment: • Lines 29-31: WTT is named the "best" predictor variable for SARs, but it is not clear that the CSS considered other inriver covariates.
• Lines 37-38: It was found here that WTT influences the smolt migration rate. But JDAR1 and Snake fish have similar migration rates. Did they have different WTT? This needs to be addressed.

Response: The actual language indicated that SARs were best described by WTT and certain ocean/climate variables. As explained in the model results, selection criteria (AICc and BIC) identified the best models, which always included the WTT variable. We agree with the reviewer that other candidate migration variables should be investigated in the future. Inspection of Figure 5.22 (old Fig. 5.21) on page 143 clearly shows that WTT between the first and third dam experienced by John Day migrants (2-5 days) was shorter than the WTT experienced by Snake River migrants (7-11 days).

BPA Comment:• Lines 42-43: SARs of downriver fish are compared to SARs from upriver fish, but these SARs are estimated over DIFFERENT reaches and distances, so we expect them to be different. The CSS needs to investigate whether the differences are more than expected.

Response: We addressed this issue above. Briefly, spawner-recruit (SR) differential mortality estimates (1.1 – 1.5) suggest about a 3-4 fold ($e^{-1.1}$ to $e^{-1.5}$) difference in life cycle survival after completion of the FCRPS. Migration distance did not change after FCRPS development; therefore, it is hard to see how different distance would drive the differential mortality response in SR. Our primary interest was whether SARs indicate the same differential mortality as was evident from the SR analyses during the post-dam period. For wild upriver/downriver SAR contrasts to date, we see a similar level of differential mortality as was evident from previous SR analyses.

BPA Comment: Page 149
• Lines 39-40: "Hydrosystem migration rates did not differ between groups but were strongly influenced by water travel time." It is not clear how to interpret this statement. Did groups have different water travel times but the same migration rate? Or did they have the same WTT? Or was migration rate and travel time examined on an individual fish basis, instead of a group basis? It is not clear.
• Lines 41-46: Distance to travel is not considered as a factor of travel time.
Page 150
• lines 1-6 – It is claimed that the "potential confounding effects due to life history differences are probably negligible," but the CSS does not attempt to model SAR using both the upstream/downstream designation and the life history differences. Additionally, the effect of distance to travel was ignored. A model that includes all possible factors affecting SAR should be considered, in order to claim that it is the hydrosystem rather than other factors that cause the difference in return rates.

Response: Sentence in question was modified to: "When Water Travel Time was incorporated in the analysis, there was no difference in migration rates between groups." The issue about distance was addressed above.

BPA Comment: In general for Chapter 5
• In order to determine if there is a biological difference that explains any differences in SAR between upriver and downriver stocks, model SAR using fork length, migration date, arrival timing, year, in addition to upstream/downstream classification. Is upstream/downstream effect significant, given presence of all others?
• Looking at population differences in fork length, migration date, etc., one at a time, is reasonable for initial data exploration, but insufficient for conclusions about the significance of the upstream/downstream effect.

Response: As discussed above, the estuary arrival timing distribution for Snake River juveniles is largely a response to the FCRPS (delay of in-river migrants, combined with a mix of project delay and barging for the transported individuals), and may not be an appropriate "independent" variable. We could pursue the remainder of the suggested analysis in future reports. However we note that SARs have been about 4-fold higher for the downriver wild populations, and none of the biological characteristics examined to date exhibit differences that would provide a plausible alternative explanation for this level of differential mortality.

BPA Comment: Throughout Chapter 5
• Typos are made in references to tables and figures throughout the entirety of Chapter 5.
• Pages 139–144: The reader is referred to a nonexistent figure for release site abbreviations.

Response: Addressed.

Chapter 6

BPA Comment: Page 154, line 9
• The notation RY has not been defined. The context suggests Return Year, but Release Year is also a possibility.

Response: MY (migration year) and RY (return year) have been defined in the final draft.

BPA Comment: Page 154, Tables 6.1 and 6.2
• Pooling migration success data across migration year and return year is valid only if those factors are nonsignificant. Perform test of homogeneity.
• Also applies to Page 155 (lines 17-23); Page 156 (Table 6.3).
Page 155, line 41
• Was return year modeled as a fixed or random effect? Most blocking factors are modeled as random effects, although there are times when a fixed effect is more appropriate.
Page 156, Table 6.3
• Chi-squared tests indicate whether there is a difference in perceived upriver adult survival across juvenile migration groups, but they do not indicate the nature of the difference. The p-values reported do not indicate that the actual ranking in the Success Rate Ranking column is significant, simply that at least one of the juvenile migration groups had a significantly different success (survival) rate than the others. One-sided tests should be performed comparing pairs of juvenile migration groups in order to test the significance of the ranking

Response: The first three bullets suggest our presentation of survival and travel time analyses (each being separate efforts) may have been somewhat confusing. As we analyzed success on a year-by-year (i.e., migration or return) and pooled (and hatchery-specific) basis using separate χ^2 tests, there was no explicit model structure for this exercise. Given pooled, MY-, and RY-specific test results, however, a formal test for year effects (a factor of secondary interest) will not change our conclusions about the principal factor of interest (outmigration experience). This is especially true given the results from our logistic regression analysis. In contrast, our GLM-based analysis of travel-time data did incorporate an explicitly defined model structure; in this exercise, return year was modeled as a fixed effect (Bullet 3).

Regarding the reviewer's last comment (Bullet 4), we presented the rankings in Table 6.3 to emphasize the consistency of ranking patterns across tests and groups. While the reviewer is correct that *post-hoc* one-sided tests could more finely resolve where the lack of homogeneity exists in the data in a purely statistical sense, this does not necessarily preclude discussion of general patterns.

BPA Comment: Page 157, Figure 6.1
• Needs error bars or confidence intervals.

Response: The estimates and CI are in Appendix D (Tables D-32 – D-36); showing the CI on the figure would result in a very cluttered graphic.

BPA Comment: Page 158, Figure 6.2
• Needs error bars or confidence intervals.

• The interpretation of Fig. 6.2, showing the proportion of LGR-detected adults and jacks detected at hatcheries, depends on the detection effort at each hatchery in each year. Without that information, it is useful only for comparing transported to inriver fish. It appears that transported fish had slightly better survival from LGR to the hatcheries, but without error bars and without information about detection effort (and harvest pressures, etc.), no real conclusion can be reached from Fig. 6.2.

Response: The 90% CI were added to figure 6.2 in the final draft. Transport and in-river CIs overlap for all years, indicating little evidence of a difference in detection probability at the hatcheries.

BPA Comment: Page 159, lines 11-15
• The overall average perceived BON-LGR adult survivals for the three migration groups are not very useful without standard errors or confidence intervals.
• It is not clear how these average survivals were computed. Were yearly estimates weighted by the number of fish returning in each year? Or were migration year estimates averaged?
• Given the finding that return year is a significant factor in perceived upriver adult survival (from the logistic regressions presented later in this chapter), pooling data over return year is not warranted.

Response: The summary that the BPA reviewer states is not useful without a presentation of confidence intervals is inaccurate, as we do present 95% confidence intervals graphically (Figure 6.4). The average "success" proportions (equivalent to the reviewer's 'perceived survival') reported on Page 159 (and plotted with 95% CIs in Figure 6.4) were computed using the pooled data (i.e., the 'Combined' field) in Table 6.1. Thus, the values presented in the figure and reported in text are unweighted averages. We also computed weighted (by return or migration year sample sizes) estimates, however, and they are virtually identical: weighted averages for Hatchery In-river, LGR, and LGS groups were 0.83, 0.76, and 0.81, unweighted values were 0.84, 0.77, and 0.81, respectively); weighted averages for Wild In-river, LGR, and LGS groups were 0.87, 0.74, and 0.90, unweighted values were 0.87, 0.76, and 0.89, respectively). Reviewers' last bullet appears incorrect. RY was significant in the travel time test, but was not in the logistic regression for adult survival.

BPA Comment: Page 160
• Lines 16-19: The model evidence ratio does not indicate that one model is "more likely" than another, in a Bayesian sense. Rather, it means that there is more evidence for one model compared to the others.
− Also applies to results for wild Chinook salmon (p. 161, lines 22-23).
− The highest evidence ratio for the best model for wild Chinook salmon (p. 161, lines 21-25; p. 162, Table 6.6) is at most 4, thus there is not clear evidence that transportation is an important factor in determining adult migration success when compared to environmental factors.
• Lines 29, 32: It is not clear how the confidence intervals on the odds ratios are computed. Provide explanation. Asymptotic normal-theory confidence intervals are considerably narrower than those reported, and do not include 1 for either LGR-transport fish or LGSdown fish. If the

confidence intervals were based on a t-distribution, the degrees of freedom should be reported (Table 6.5).

Response: The reviewer's first comments are a matter of semantics, not a technical or analytical issue necessarily. We used these model fit criteria as one (among others provided) to judge which model(s) best explained the observed data. The reviewer is taking literary license with what we said in the text, as our conclusion based on model results was that there is stronger support for a transportation-legacy hypothesis than an environmental conditions-only hypothesis. The confidence intervals shown on page 160 for the odds ratio of parameter **LGR** relative to parameter **In-river** and parameter **LGS down** relative to parameter **In-river** are obtained from running a binomial logit in SYSTAT (logistic regression) with a categorical variable **transport** (split on three levels: In-river = 0, LGR = 1, and LGSdown = 2) and the other non-categorical variables modeled. An exponential transformation of the logistic regression parameter estimates for LGR and LGSdown will provide the odds ratio of these parameters relative to In-river. SYSTAT prints out the odds ratios and 95% confidence intervals directly, but they may also be obtained by taking exponents of the logistic regression parameter estimates of LGR and LGSdown and their 95% confidence intervals. Table 6.5 shows a logistic regression parameter estimate and standard error for LGR of -0.446 and 0.092, respectively, from which the 95% CI is $-0.446 \pm 1.96 \cdot 0.092 \Rightarrow (-0.6263, -0.2657)$. The exponential transformation results in a odds ratio of 0.64 and 95% confidence interval of (0.53, 0.77) as shown on page 160. Likewise, Table 6.5 shows a parameter estimate and standard error for LGSdown of -0.212 and 0.123, respectively, from which the 95% CI is $-0.212 \pm 1.96 \cdot 0.123 \Rightarrow (-0.4531, 0.0291)$. The exponential transformation results in an odds ratio of 0.81 and 95% confidence interval of (0.64, 1.03). It is not clear how the BPA reviewer computed narrower asymptotic normal confidence intervals for the odds ratio.

BPA Comment: Page 161
• Table 6.5
− Degrees of freedom should be reported for each parameter estimate.
− Surprisingly, warmer temperatures were associated with higher perceived adult survival. Perhaps temperature is confounded with run (spring versus summer).
• Lines 29-30: The odds ratio is misinterpreted here. An odds ratio of 0.5 does not mean that the probability of success of LGR-transport fish is half that of inriver fish. If the probability of success (i.e., perceived adult survival from BON to LGR) is for LGR-transport fish, and is for inriver fish, then:

$$[\text{Odds ratio} = ½, \text{ then } P_{LGR} = P_{inriver}/(2-P_{inriver})]$$

This means that the probability of success of LGR-transport fish depends on the value of the success probability for inriver fish, as demonstrated in Table 2 below. Table 2 indicates that for an odds ratio of 0.5, the probability of success of LGR-transport fish is generally greater than half that of inriver fish, except for very small inriver success probabilities, which are not applicable here.

Response: We changed the language on p. 161, lines 28-30 to more accurately reflect interpretation of the odds ratio as follows: "Further, the odds ratio estimate for the LGR group

(estimate: 0.46; 95% CI: 0.26-0.84) indicates that these adults had significantly lower odds of surviving their BON-LGR migration than in-river outmigrants (i.e., the 95% CI did not include 1)."

BPA Comment: Pages 162-163, Hatchery Chinook arrival and travel time ANOVAs
• For both arrival time and travel time, the interaction term between return year and juvenile migration (outmigration) method was significant. This affects interpretation of the main effects of both return year and outmigration method, so conclusions based on the main effects alone are invalid.
• The ANOVAs should be included in the report.

Response: The reviewers are mistaken in implying we drew conclusions about main effects on arrival time and travel time. We accurately reported the results of the interactions.

BPA Comment: Page 164, lines 38-40
• How much of TIR or D is explained by observed differences in perceived upriver adult survival between inriver and LGR-transport fish?

Response: In the conclusions, we were simply noting that a portion of deviation in *TIR* and *D* may be attributable to survival differences occurring in the mainstem after adults return .We did not attempt to quantify this phenomenon. Based on future priorities, this could be a focus for future studies.

Chapter 7

Page 168 (lines 31 and 46) – two comments regarding input values to the simulator program.

> i). Survival from release to LGR = 0.95 seems high, and does not correspond to year 2000 data used as basis for default values.
>
> Response: Migration year 2000 data was used to establish the default curves for survival rates, collection efficiency at dams, inter-dam travel times, and the initial arrival timing distribution at LGR. The survival from release to LGR was simply set at 95% to reflect a typical survival rate from the head of the hydrosystem at Lewiston to LGR. The release size of 32,000 fish was aimed at providing an arrival population at LGR of approximately 30,000 fish, which is in the range observed with wild Chinook as well as hatchery Chinook from Rapid River and McCall hatcheries in several years. Since program computing time increases somewhat exponentially as release number increased, a higher release number and lower survival rate from release to LGR to achieve approximately 30,000 fish arriving LGR would have increased the overall computing time without affecting the simulation outcome.
>
> ii). An SAR=0.03 seems high, given that observed SAR has been lower than the target value of 2% in most years according to this report.

Response: Only one SAR level was simulated for this report, with those fish having capture histories reflective of a particular study category getting assigned an adult return based on the random binomial draws from the number of smolts in that particular study category prior to any expansion to LGR equivalents. Since the assumptions being tested in the 12 scenarios run to date related to temporal changes in inter-dam reach survival rates and collection efficiencies at dams, and not to temporal changes in SARs based on timing of smolt arrival at LGR, we did not need to run more than one level of SAR. Had we set the input SAR at 2% or 1%, our ability to investigate biases caused by violation of the CJS model assumption that "all fish in a release group have equal detection and survival probabilities within the same river reach or at the same dam" (Assumption #2 in Appendix C) would not have been affected. The resulting population variability about the SARs, *TIR*s, and *D* would increase as the input SARs level got smaller, but this effect is unrelated to the CJS model assumption being tested.

Page 169 (line 9) BPA Comment: comment that the joint probability of survival from BON to TWX and detection at TWX =0.10 is high based on past years.

Response: A lower joint probability could have been used, but it would not have affected our evaluations of impacts of violations of Assumption #2 described in the previous response. We allowed temporal changes to occur in reaches and dams between LGR and MCN, and maintained the same default inputs for all reaches and dams below MCN as well as at the trawl in the 12 scenarios tested. A lower joint probability assigned at the TWX would have reduced the number of smolts caught in the trawl and thus increased the population variability for the S_R and D parameters to some extent, but again as in the previous response, this effect is unrelated to the CJS model assumption being tested.

Page 170 (lines 9-11, 21-23, and 39-41) and page 171 (lines 8-10) – comment that survival probabilities used in simulation scenarios #5, 7, 10, and 12 include inriver survival probabilities >1, when the variable day is 0 or very low. Inriver survival should be parameterized using only admissible parameter values (i.e., ≤1) and included in this report.

Response: In the simulator program, we have constraints on the daily values taken from the parabolas and linear trends to avoid the problems the reviewer expressed. The survival rate and collection probabilities are not allowed to exceed 0.95 or drop below 0.05, in order to keep the random beta distribution draws from occasionally trying to return an undefined value (>1 or <0), which terminates the run. Figures 7.8 and 7.9 show the how this constraint changes the steepest linear trends evaluated to flat lines before or after certain dates. In the methods section where the trend lines for the 12 simulation scenarios are presented, we will add text to indicate that daily values taken from the parabolas or linear trends are constrained between 0.05 and 0.95 prior to these values being used in beta draws for survival rate and collection probabilities that are finally used in the binomial draws for numbers of fish surviving as well as collected each day within the various inter-dam reaches and dams.

Page 171 (last paragraph) and page 174 – comment regarding SIM-2 where the emphasis on the T and R groups is confusing. The review comment goes on to suggest that a simpler method of assessing the effect of detection-influenced survival would be to simulate date under the

scenarios described (post-turbine survival < post-bypass survival < post-spill survival, with varying proportions of undetected fish passing via turbine or spill) and examine estimates of C_0, C_1, and T_0.

Response: The simulator program was designed to address the impacts on CJS estimates of survival rates and collection probabilities when the underlying "true" survival rates and collection probabilities are changing temporally. This condition causes violation of the CJS model assumption that "all fish in a release group have equal detection and survival probabilities within the same river reach or at the same dam." It was not designed to address the impacts caused when prior detection history causes a change in later downstream survival rates and collection efficiencies. With the start in 2006 of pre-assigning PIT-tagged fish into Group T which reflects the experience of the run-at-large (untagged and tagged) and Group R which is used for estimating in-river survival rates. The attempt in the draft report to address the potential impacts of prior detection history in the indirect approach utilized was determined by the CSS Oversight Committee to be too ambitious given the tight deadlines for the 10-year report. Therefore, we have deleted SIM-2 from the report.

Page 175 (last paragraph) – comment states that it is not clear if the "true" survival parameters used to compute LGR equivalents are averages of seasonal survival parameters, or if LGR equivalents are computed on a daily basis and then summed over the season. Give the temporal variation in survival parameters introduced in these simulations, the latter approach should give a better representation of the "true" C_0, C_1, and T_0 groups. Clarify approach and, if necessary, rerun simulations.

Response: The known (i.e., "true") S_2 and S_3 used to convert smolt counts to LGR equivalents are obtained in three steps: 1) survival rate from LGR to LGS is obtained by dividing the LGR computed "known" number of fish remaining inriver (after subtracting off the removals for transportation) for the season into the LGS "known" number of fish surviving there, which is computed by summing over the season the daily number of fish assigned as survivors based on binomial draws each day with survival rates obtained from the daily trend relation (parabola or linear); 2) the travel time from LGR to LGS distributions will shift the surviving fish at LGS into their starting dates there; 3) survival rate from LGS to LMN is obtained in the same manner as step 1 (simply substitute LGS for LGR and LMN for LGS). When step 3 is completed the "known" number of fish surviving to LMN is obtained. This process has produced S_2 and S_3 that are based on total "known" fish arriving the downstream dam divided by total "known" fish continuing inriver from the upstream dam. This process produces the proper "known" parameters S_2 and S_3 for use in converting downstream smolt counts into LGR equivalents. The approach preferred by the reviewer would be much more difficult to implement, but should provide the same starting population at LGR, if done correctly.

Page 194 – comment is split into two parts due to length followed by answers.
Comment – The CSS uses results of the second set of simulations to address how to best analyze data using the NPT approach, in which tagged fish are pre-assigned in to migration groups: T (transport) fish are transported upon their first detection at a transport dam; R (river) fish are returned to river upon all detections. Using the $C_0/C_1/T_0$ approach to analyze data with pre-assigned migration groups is not intuitive.

Response: This is incorrect as stated. The goal of the second set of simulations (now dropped as too preliminary) was not to address how to best analyze data using the NPT approach, which has been implemented in the CSS starting with the 2006 migration year. Rather it was aimed at showing how the categories C_0, C_1, and T_0 utilized in CSS analyses may best be computed. In the current CSS, annual estimates of overall SAR must be computed as a weighted combination the category-specific SARs, where the weights are the proportion of the run-at-large (untagged and tagged) represented by each category. Using the NPT approach, Group T will provided the annual estimate of overall SAR directly. The reviewer also mischaracterized the fish in Group T as being transported upon their first detection at a transport dam. At collector dams, Group T fish go the direction of the untagged fish, regardless of whether that is to raceways for transport or back to the river. Likewise, if untagged fish are being transported from a dam, then any fish in Group T detected at that dam will also be transported, regardless of whether that fish had been previously detected at dam upstream.

Comment: It would be simpler and more defensible to simply compare the SAR of the T group to the SAR of the R group. All "R" fish will have migrated wholly inriver, while some "T" fish will have been transported and others (undetected) will have migrated inriver. The comparison of SAR(T) to SAR(R) is more easily interpreted for management, because the alternative to transportation is to return detected fish to the river, whereas the transportation alternative being tested in the SAR(T0) vs. SAR(C0) comparison is not clear.

Response: When analyzing the data collected from migration years 2006 and later, we will be comparing SAR(T) to SAR(R) as the reviewer suggests, but this does not preclude the utility of additional comparisons among all three study categories C_0, C_1, and T_0. Just as we have a time series of SAR(T_0), SAR(C_0), and SAR(C_1), and overall $SAR_{LGR\text{-}to\text{-}LGR}$ (akin to SAR(T)), we will also be able to compare SAR(R) data with prior years by substituting SAR(C_1) for SAR(T_0) in the formula of overall $SAR_{LGR\text{-}to\text{-}LGR}$ for pre-2006 migration years. The reviewer failed to include the fact that in addition to fish transported and those undetected at collector dams, Group T may include fish bypassed at Snake River collector dams during April and early May under the policy begun in 2006 of delaying the start of transportation at those dams.

Chapter 8

Page 198 Lines 35-38: The trend of performance measures for wild fish mimicked the overall trend of performance measures for the collection of hatcheries, but did not agree well with the trend from any single hatchery across all years. It is not clear which single hatchery could be used to make inference to wild fish. Also applies to Pages 199-200, bullet (b) of Chapter 5 summary.

Response: Hatchery Chinook salmon and wild Chinook salmon responded nearly identically to environmental and/or seasonal conditions in terms of their fish travel time, instantaneous mortality rates, and survival rates in the LGR-MCN reach. Thus, hatchery Chinook salmon provide valuable information on the response of wild Chinook salmon to conditions experienced in the hydrosystem.

Differential mortality between upriver and downriver stream-type Chinook populations has been estimated for wild populations from both spawner-recruit (Schaller et al. 1999; Deriso et al. 2001; Schaller and Petrosky 2007) and PIT-tag SAR (CSS study) data sources. The CSS also investigated whether a similar level of differential mortality was present between PIT-tag SARs for five upriver and one downriver hatchery Chinook populations. Because biological characteristics of a population could differentially influence survival to adult return (see above), we also summarized hatchery pre-smolt FL at the time of tagging, and hatchery smolt arrival timing distributions entering the hydrosystem (LGR or BON) and arriving at the estuary (BON).

Upriver and downriver hatchery spring/summer Chinook SARs did not show the same level of differential mortality as was apparent from the wild populations. Survival of hatchery fish is subject to additional fitness and rearing factors that may not affect wild populations. CSS currently has the ability to compare SARs from a single downriver hatchery (Carson NFH) with those from five Snake River hatcheries. 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.

Although Snake River hatchery Chinook exhibited a generally more positive response to transportation and relatively lower levels of differential mortality than wild populations, annual SARs of wild and hatchery Snake River Chinook were highly correlated. In view of this high correlation, continuing the CSS time series of hatchery SARs will be important to augment wild Chinook SAR information following future years of low escapements, in addition to providing valuable management information for the specific hatcheries. One advantage of the CSS study is that tagging takes place at the hatcheries and in the tributaries for wild populations. This approach allows for detecting different responses to management actions for different components of the wild and hatchery aggregate groups, unlike approaches that only tag at the upper most dam. Finally, it is of interest to the region of how the specific hatchery groups respond to the hydrosystem management actions. The reviewers suggest a much smaller number of PIT-tagged hatchery fish could be used. We believe that the sample sizes should be periodically reviewed based on updated survival estimates, and regional monitoring and evaluation needs.

Page 198, Report confidence intervals for results (e.g., geometric means).

<u>Response</u>: Confidence interval results are presented in Appendix D.

Page 199 • Lines 32-35: The inference made from declining SAR(C_1) over the season to hydrosystem-caused post-Bonneville mortality is unfounded. There are alternative possible causes of post-Bonneville mortality, including temperature, pollution, disease, and seasonal changes in estuary conditions. No conclusions about the relative importance of the various potential sources of mortality can be reached here.

• Lines 40-42: The CSS claims that Snake River wild steelhead SARs averaged less than 2%. It is difficult to confirm this statement, because the annual SARs are not presented in tabular form in this report. However, Fig. 3.12 suggests that average annual SAR for wild steelhead may be greater than 2%. Document annual SARs in the table and explain apparent inconsistency.

Response: No unambiguous demonstration of the effect is claimed; the report states that the declining SAR is "consistent with the hypothesis" of protracted migration-induced mortality. Most of the commenter's listed alternative causes are actually mechanisms which could cause mortality due to protracted migration. For instance, temperatures increase over the season (for spring migrants). Disease expression can be affected by protracted migration, through delaying of saltwater entry relative to smoltification and through exposure to higher temperatures. Seasonal changes in the estuary are another likely candidate for mortality induced by late arrival of smolts. If the commenter has evidence that seasonal distribution of pollutants in the estuary can explain such a dramatic drop in post-Bonneville survival over the season, we would be eager to see it.

Evidence for wild steelhead SARs averaging less than 2% can be found in Table D-19, where transport SAR averages slightly over 2%, but in-river SARs average less than 1%. Annual overall steelhead and Chinook SARs are also found in Appendix E, which will be included in the next draft of the report. Further, the Chapter 4 weighted mean wild steelhead overall SAR is 1.95% (Figure 4.4).

Page 200, Lines 8-14: The CSS did not compare the ratio of upstream and downstream SARs to in-river survival between Lower Granite and John Day, so the conclusion that upstream fish experience extra mortality caused by the hydrosystem is unjustified. Also applies to Page 200 (lines 33-34).

Response: The reviewers' comments contain a purported comparison of survival from John Day to Bonneville Dam with the ratio of SARs from upriver and downriver stocks (Table 1), and assert that this comparison would be more appropriate than a SAR comparison that indexes smolts leaving the production areas (i.e., at the first dam). There are two problems with the reviewers' approach. Their proposed approach is inconsistent with the original SR definition of differential mortality (e.g., Deriso et al. 2001), where spawners were indexed at the spawning grounds and recruits were indexed at the Columbia River mouth (p. 116, lines 29-31). Second, the reviewers propose to account only for the passage mortality experienced by in-river migrants and not that of transported smolts (the migratory route the majority of fish experience). One could, in theory, fix the smolt indexing location at any number of locations (JDA or BON), but this would be a very different analysis, and not consistent with the SR based estimates of differential mortality. It is not clear what the reviewers' proposed adjustment only for in-river survival would accomplish, other than further confuse this issue.

Page 200, Lines 23-24: The claim is made that that the CSS shows clear evidence of delayed estuary entry of Snake River in-river smolts, caused by passage through the hydrosystem, on the basis of comparisons with John Day smolts. This is not true. The CSS found that Snake River and John Day smolts (1) initiate migration at the same times, and (2) migrate at similar rates through the first three dams passed. Given the extra distance traveled by the Snake River smolts,

it is not surprising that Snake River smolts enter the estuary later than John Day smolts. The CSS analysis would be more useful if it had compared the observed and expected arrival dates of the Snake River fish, given their migration initiation date, migration rate (through the first three dams), and distance to travel.

Response: The BPA reviewers appear to be confused on the purpose of the upriver/downriver analysis, which was stated in the Chapter 5 introduction (p. 106): "*our specific interest ... is whether upriver/downriver differences in SARs for wild and hatchery stream-type Chinook were consistent with the differential mortality estimated from SR [spawner-recruit] models for wild populations"* Previous published SR analyses indicated there was a systematic increase in mortality for Snake River populations, which did not occur in the downriver populations, associated with the construction and operation of the FCRPS (e.g., Deriso et al. 2001; Schaller et al. 1999). In the SR model formulations, any differences in smolt mortality caused by different travel distances would be incorporated into the intrinsic productivity (Ricker "a"). Obviously, the migration distance for upriver and downriver populations did not change over the time period of FCRPS development; Water Travel Time (WTT), Fish Travel Time and hydro impacts did change with this development. WTT for Snake River stocks before FCRPS development were only about 2-3 days; Snake River smolts were historically able to arrive at the estuary more in synchrony with their morphological, physiological and behavioral development (e.g., Budy et al. 2002; ISG 1999). Available evidence from a mostly free-flowing migration corridor (Whitebird trap on the Salmon River to Ice Harbor Dam) also suggests smolt survival was high before FCRPS development (Raymond 1979). Applying the survival per mile from the Raymond study the information suggests that the historic survival from Lewiston to Bonneville dam was over 90%.

Page 200, Lines 26-30: The conclusion that differing seasonal SARs for upstream versus downstream smolts is evidence of delayed mortality ignores possible alternative explanations, including potentially different ocean residencies.

Response: Based on the weight of evidence in the peer-reviewed literature, it is apparent that the highest level of mortality takes place in the first year of ocean residence.

Page 200, Lines 37-42: It appears here that wild and hatchery Chinook salmon transported from LGR always had 10% lower SAR than fish passing through the hydrosystem by alternative routes. It should be noted that the effect for hatchery fish (4% to 7%) was considerably less than the effect for wild fish (15%), so the 10% effect reported is somewhat misleading.

Response: In this comment, the reviewer has confused adult upstream survival rates with SARs.

Page 202, Lines 11-16; lines 39-41: The claim is made that the CSS addresses the question of whether smolt transportation compensates for effects of the Federal Columbia Power System (FCRPS) on survival of Snake River Chinook and steelhead. This claim extrapolates past the available data. The CSS compares the SAR of transport fish to the SAR of fish migrating in-river. While the in-river fish experience effects of migrating through the FCRPS, available data do not

indicate the magnitude of those effects; this would require comparing the SAR of fish migrating through the FCRPS to the SAR of fish migrating through the same reaches but not through the FCRPS. That is not possible. At most, the comparison of the SAR of transport fish to the SAR of in-river fish indicates whether transportation is a viable management option; it is not equivalent to comparing transportation to migration through the unimpounded river. It is worth noting that the SAR from BON to BOA for hatchery Chinook salmon from the John Day river was less than 2% for 2001 through 2004 (Table 5.10). Regardless of the validity of upstream-downstream comparisons, these low SARs for John Day fish suggest that the hydrosystem is not the only factor in below-target SARs.

Response: The reviewer has misconstrued the analyses conducted within the CSS. The CSS has monitored the effectiveness of transportation versus in-river migration in the presence of the FCRPS. We have also evaluated those SARs relative to the NPCC's 2-6% SAR objectives. We make no statements regarding survival in an unimpounded river. The reviewer makes references to hatchery Chinook salmon from the John Day River, which do not exist. It is important to note that the wild Chinook SAR from the John Day River has met the NPCC SAR objectives, providing evidence that stocks which migrate through fewer dams can meet these interim survival objectives.

Page 202, 3rd paragraph

Response: The geometric mean is a better measure than arithmetic mean of central tendency for right skewed (log-normally distributed) distributions such as *TIR* and *D*. They both represent ratios of survival rates, for which the ordering (i.e. which is numerator and which denominator) is arbitrary. From Zar (1984, p. 24): "[The geometric mean] finds use in averaging ratios where it is desired to give each ratio equal weight".

The wording about steelhead *D* will be changed to indicate the evidence about whether *D* is in general less than 1 for wild steelhead is ambiguous. The implications of *D* being less than one while *TIR* is greater than one will be noted. The question of whether or not to transport depends in large part on what the alternative to transportation is. The value of *TIR* serves to answer this question in some contexts, but not in others. If the only alternative is simply to allow migration in-river under current configuration and operation, *TIR* is a useful metric. If the range of alternatives included strategies to significantly improve in-river migration conditions, up to and including dam breaching, then *D* tells us more about any expected benefits that might be derived from these alternative strategies.

Page 202, Last paragraph: The CSS compares observed SAR estimates from PIT-tagged fish to the NPCC objectives for SAR (2% minimum, 4% average), without addressing the NOAA finding that PIT-tagged fish have lower survival than untagged fish (as requested by the ISAB). Without knowing the size of the PIT-tag bias, comparisons of PIT-tag SAR to target values are not completely useful.

Response: The introduction to Chapter 5 (p. 105) cites the ISAB (2006) issue that more attention should be given to whether PIT-tagged fish survive as well as untagged fish. Chapter 5 contains a section (p. 147) titled: "Do PIT-tag SARs represent SARs of the run-at-large?" with further

discussion on p. 150-151. We agree with the ISAB (2006) conclusion that more attention should be given by CSS *and the Region as a whole* (emphasis added) to the discrepancy of SARs between PIT-tagged and untagged fish. However, the extremely tight reporting requirements did not allow for an examination of all the assumptions and data adjustments currently necessary to estimate SARs of the untagged component. Because the issue involves potential bias of both run-reconstruction and PIT-tag methodologies, resolution will require a collaborative effort among several technical groups in addition to the CSS project.

Contrary to the NWFSC comment that no caveat exists that PIT-tagged SARs may have a bias relative to the NPCC goal, the draft report explicitly stated (p. 147) "[t]he primary concern of negative bias from PIT-tag SARs would be in evaluating whether SARs are meeting NPCC biological objectives (2% minimum, 4% average)." Also, "[i]mplications of bias (if present) would be negligible for relative comparisons of the CSS PIT-tag SAR data, such as between Snake River migrants with different hydrosystem experiences, or between Snake River and downriver populations." We also point to future monitoring and evaluation tasks to help resolve this issue in the future. We note that the 2 to 4 % goal itself was based on analyzes involving tagged fish that presumably experienced some handling mortality relative to the unmarked population.

Page 203, 3rd paragraph

Response: We agree that we have not performed a "comprehensive" analysis of strategies for varying transportation over the season, and we don't believe we implied that. The CSS was not designed primarily for that purpose. However, we have explored seasonal variation in reach survival and transport and in-river SARs and found some interesting results, and we believe that "[Results] have the potential to inform management on when to initiate transportation" is cautiously and appropriately worded.

The C_1 group is the appropriate group of interest for comparison to transported fish for some management questions, and we used this group in the seasonally varying SAR estimates. For instance, if the question is simply "if a fish is collected, then given when it is collected, should it be transported?", this group is appropriate. However, the question of when to turn on or off transportation says nothing about the alternative to transportation, i.e., how the river would be managed for spring migrants in the absence of transportation. Depending on management actions (e.g., high spill at collector projects), there could be a large percentage of C_0 fish in many years. Then, the question is, "When is transport SAR greater than in-river SAR, given that in-river fish would be some mix of C_0 and C_1 fish?" The appropriate weighting of the two in-river SARs would depend on the proportions in each group expected under the particular management regime.

Appendix B

Page B-3, Figure 1 – BPA Comments:

- *The estimators of $\acute{\emptyset}_1$, $\acute{\emptyset}_2$, and $\acute{\emptyset}_3$ are correct.*

- *The figure is somewhat cryptic. The parameters $\acute{\emptyset}_i$ are not defined, nor are the statistics R_i, $R'_{1\cdot 2}$, $R'_{12\cdot 3}$, …, r_i, m_i. The reduced m-matrix is not so standard that the CSS should expect all readers to recognize and understand it without further explanation. Provide more detail.*

Response: In order to help the reader understand the notation in Figure 1, we expanded the text to include a detailed description of all notation and concepts being illustrated in Figure 1. The reduced *m*-matrix (detailed in the Burnham et al (1987) monograph) is a useful summarization of all data required to estimate the parameters of inter-dam reach survival rates ($\acute{\emptyset}_i$) and dam collection probabilities (p_i). It should be familiar to those who have used the CJS model. For those unfamiliar with the CJS model, the schematic with legend should help them better understand the estimation process.

Page B-4 – BPA Comments:

- *The CSS explains that they allow individual reach survival estimates exceeding 100% when computing an overall multi-reach survival estimates. Why, then, do they not allow $S_{JDA\text{-}BON} > 1$ for 2004 for Carson NFH Chinook in Chapter 5?*

Response: In Chapter 5, the CJS based estimate of survival from release at the hatchery to Bonneville Dam was >1 for the Carson NFH Chinook in 2004, not a survival between JDA and BON as stated by the reviewer. In that situation, we felt an average release-to-BON survival rate of the prior years would be better estimate than simply constraining the estimate to 1. This was the first occurrence of a release-to-first dam estimate of survival exceeding 1. Between adjacent reaches, the CJS estimates of survival have an inherent negative correlation, since the estimated population in the tailrace the upper reach becomes the starting population in the next reach downstream. When one estimate is high, the next will be low, and visa versa, as one travels down through all reaches. Therefore, when we take the product of a series of reach estimates to obtain a longer multi-reach survival rate, the reach-to-reach variation is dampened in these longer reaches, thus balancing the effect of some individual estimates being >1. A greater concern is having individual reach estimates of very poor precision lower in the hydrosystem due few fish there. Therefore, we would not used an estimate with CV >25%, and would extrapolate the survival of that reach based on a per-mile survival rate based on the available upriver multi-reach survival rate estimate.

- *The verbal description of the weighted average of survival estimates provided in the second full paragraph is insufficient. An equation demonstrating precisely how the overall survival estimates was estimated is required.*

Response: The distribution of PIT-tagged fish detected at LGR is partitioned into strata. The program allows strata defined by equal proportion of fish per strata or equal number of days per strata. The CJS is run separately fish in each strata and then common reaches across strata are weighted by inverse relative variance times proportion of run-at-large (untagged and tagged fish) for wild Chinook and simply inverse relative variance for Chinook from each hatchery. This approach was only used on Chinook in the early years of the CSS, prior

to any analyses on steelhead. Details of the computation of the weighted average survival rate in the jth reach are as follows and has been added to Appendix B:

1. Let A_k = proportion of annual passage index data from Smolt Monitoring Program in the kth stratum
2. Let B_k = theoretical variance of CJS estimates for k^{th} stratum,
   where $B_k = S_{jk}^2 [1/r_j - 1/R_j$ + additional terms shown for var($\acute{\emptyset}_j$)
   on Page 115 of Burnham et al. (1987)]
3. Let S_{jk} = estimated survival rate of jth reach in k^{th} stratum
4. Weight for wild Chinook is $W_{1k} = (A_k)(S_{jk}^2/B_k)$ in k^{th} stratum
5. Weight for hatchery Chinook is $W_{2k} = (S_{jk}^2/B_k)$ in k^{th} stratum
6. Weighted estimate across k strata for jth reach is:
   $\Sigma (W_{1k})(S_{jk}) / \Sigma (W_{1k})$ for wild Chinook
   $\Sigma (W_{2k})(S_{jk}) / \Sigma (W_{2k})$ for wild Chinook

- *The CSS used weights equal to the inverse relative variance of the reach- and cohort-specific survival estimates. How were the variances of those reach and cohort survival estimates computed? How was the standard error on the weighted average survival computed? Provide details.*

Response: In the sub-cohort approach to estimating reach survival rates, the fish detected at LGR were stratified into a user defined number of strata (a sub-cohort is simply a stratum). The standard CJS model is used separately with those fish re-released at LGR in each stratum. Once the CJS estimates of survival are obtained, the standard theoretical variances of the CJS model, in the form of inverse relative variances as shown in the previous response are used to weight each stratum's survival rate for a particular reach, and summed to create the weighted average reach survival rates for that particular reach. The reviewer should note that the sub-cohort approach was not used in the 10-yr report. All estimates of reach survival rates are based on the CJS model applied to the full sample of fish released, rather than simply on those detected in temporal intervals at LGR.

- *In the final partial paragraph, the CSS discusses using a "per-mile" expansion of juvenile survival in cases where it was impossible to estimated survival to BON directly. Previously (Chapter 3), they used a per-km method of extrapolation. Either there or here, did they consideration other basis for extrapolation? Did they consider the goodness-of-fit of the extrapolation method used? Did they estimate the standard error on the survival estimate to BON, either with or without the extrapolation?*

Response: In the bootstrap computer program that computes all parameter estimates along with the confidence intervals, both a "per-mile" and "per-project" extrapolation is computed. The reference to "per-km" extrapolation in Chapter 3 will be revised to "per-mile" extrapolation. The rationale for choosing the "per-mile" extrapolation approach as the standard instead of the "per-project" approach has been detailed in a prior response in Chapter 3 to the same BPA comment. Goodness-of-fit was not computed. Bootstrap standard errors and confidence intervals are computed in the bootstrap computer program, and will be added to the appropriate Appendix D tables as stated in a prior response to a BPA

comment on Appendix D. In the cases were an extrapolation was necessary to in order to obtain an estimate to BON, the concept of estimating a standard error on the survival rate without the extrapolation as suggested by the BPA reviewer does not make sense.

Page B-5 – BPA comment:

- *The CSS lists the three ways in which fish can pass an individual transport dam, and indicates that these three passage routes describe the passage routes through the hydrosystem. However, their three passage routes must be combined over multiple dams to describe the possible passage routes through the entire hydrosystem. For example, there are seven possible passage routes through LGR, LGS, and LMN that result in transportation from one of those dams – (i) transport at LGR (route 1), (ii) transport at LGS following either detection at LGR, or non-detection at LGS (routes 2 and 3), and (iii) transportation at LMN following either detection at both LGR and LGS, detection at only one of LGR and LGS, or non-detection at both LGR and LGS (routes 4-7). Thus, the CSS "partition" of PIT-tagged smolts arriving at LGR is, at best, unclear form their description and, at worst, potentially omitting considerable numbers of fish. Clarification in this report is required.*

Response: The CSS does not attempt to analyze all "possible" routes of passage in the manner inferred by the BPA reviewer. Instead, the CSS has created the three groupings of "possible" passage routes that best reflect what is being experienced by the untagged run-at-large. For the migration years covered in this CSS 10-yr report, the untagged run-at-large was most often transported at the three Snake River collector dams if collected there (exception is 1997 when management operations bypassed many tagged and untagged fish at LGS and LMN during parts of the migration season). We say that the collected fish were most often transported rather than 100% transported, since there are occasions over the years when all fish from raceways were returned-to-river due to unavailability of enough barges at peak passage times, or malfunctions at the facility that required short-term bypassing of all fish. Given the project operations from 1994 to 2004, the untagged run-at-large was either (i) collected and transported from one of the three Snake River transport site, (ii) collected and bypassed from one or more of these sites, or (iii) uncollected at these three sites, passing through either spill or turbines. For transported fish, the CSS utilizes those either transported from LGR, or first-time detected fish that are transported at LGS or LMN. We rely on first-time detected PIT-tagged fish at the two downstream dams, since those PIT-tagged fish match closest to the untagged run-at-large. Since we must return fish from the collector dams each year in order to estimate the inriver reach survival rates, there are occasions when these fish will be collected at the downstream sites and transported. Generally, all fish subsampled and handled in the Smolt Monitoring Program at these dams will go to transportation after handling and recovery. However, most multi-site detected PIT-tagged fish that get transported do not reflect the untagged run-at-large. Therefore, the BPA reviewer's contention that the CSS is "potentially omitting large numbers of fish" is incorrect.

Page B-6 – BPA Comment refers to "#5 Observed transportation estimate of run-at-large smolts at LGR is t_2 = (LGR run-at-large transported/LGR run-at-large collected) m_{12} and expectation of

$E(t_2) = E(m_{12})\ P_{t2}$ where P_{t2} is the proportion of run-at-large (total fish at level of species and rearing type from Smolt Monitoring Program) transported at LGR".

- *#5. Is "run-at-large" equal to "untagged" here, or does it also include tagged fish.*

Response: The numbers of run-at-large fish collected and transported at LGR include both untagged and tagged fish. The Smolt Monitoring Program provides separate estimates of collected and transported "unclipped, non-CWT" yearling Chinook, which we use for run-at-large wild Chinook, "clipped or unclipped with CWT" yearling Chinook, which we use for run-at-large hatchery Chinook estimates, "unclipped" steelhead, which we use for wild steelhead, and "clipped" steelhead, which we use for hatchery steelhead.

- *How is P_{t2} estimated?*

Response: This parameter is an estimate of the proportion of PIT-tagged fish that would have been transported at LGR if PIT-tagged fish had been transported at the same rate as the run-at-large (see prior response for definition of run-at-large fish). It is estimated as (est. run-at-large transported)/(est. run-at-large collected) for the group of fish of interest.

- *Is P_{t2} really the proportion of the entire run-at-large that were transported at LGR, or only the proportion of the run-at-large collected at LGR that were transported?*

Response: P_{t2} = (est. run-at-large transported)/(est. run-at-large collected); therefore, it is the proportion of the run-at-large collected at LGR that were transported. We multiply P_{t2} with m_{12} to get t_2.

- *Similar comments pertain to #7 and #9.*

Answer: The same response for LGR (#5) applies to LGS (#7) and LMN (#9).

Page B-7 – BPA Comments:

- *#13 - #15: It is essential for the CSS to actually write out the expected values of the statistics T_0, T_0*, and C_1 in terms of the underlying model (i.e., survival, detection, transportation, and removal parameters), rather than leaving them partially defined. This level of technical detail is essential for all readers to know exactly what is being estimated by the parameters in the report.*

Response: The details requested by the BPA reviewer already exist in #1 to #12. In order to simplify the long formulas for the expectations of T_0, T_0*, and C_1, we feel our presentation is actually easier for readers to visualize what is being estimated. See Appendix C for formulas of expectation for T_0.

- *#15, #16: The statistics d_0 and d_1 are never defined. The 50% survival probability is not explained – 50% survival to where? On what basis is 50% chosen? Why not use the actual estimated survival probability to whatever site or sites are used?*

Response: The parameters d_0 and d_1 are defined directly below the formula and their rationale detailed in the first full paragraph on page B-8. These parameters account for PIT-tagged fish from categories C_0 and C_1, respectively, removed below LMN. Since most of this type of removal occurred at MCN in 1994 and at JDA or BON in other years, and survival from LGR to these sites was approximately 50% in the years affected, we developed into the bootstrap program a fixed 50% removal adjustment for all years. Although a year-specific estimated removal rate could have been programmed, we opted for this simpler approach when programming for this adjustment since the numbers of PIT-tagged fish affected was relatively low (numbers are presented in response to the next BPA comment).

This same basic question was raised by BPA in their review of the 2006 CSS Annual Report, and our response to them then (Berggren et al. 2006, pages 165-166) is still pertinent. "PIT-tagged fish not confirmed as being returned-to-river at a downstream dam needed to be removed from either the C_0 or C_1 study groups. Fish were considered as removals at McNary Dam when detected on the raceway or sample room monitors or only on the separator monitor during the summer transportation season, or when collected and removed at John Day or Bonneville Dam for other research purposes. Samples of CSS PIT-tag hatchery Chinook from Rapid River, McCall, and Dworshak hatcheries were collected and sacrificed at John Day and/or Bonneville dams during migration years 1999 to 2003 for physiological (blood chemistry) evaluation (Dr. Congleton, University of Idaho Fish and Wildlife Unit). Because most removals occurred at John Day and Bonneville dams for other research purposes, we settled on a fixed 50% Lower Granite to Bonneville Dam survival rate for each removed fish in order to subtract these fish in LGR-equivalents from the estimated number of smolts in Categories C_0 and C_1. Most survival rates from Lower Granite Dam to Bonneville Dam from 1995 to 2004 (excluding 2001 when extremely low in-river reach survival rates occurred on in-river migrants) have been averaging around 50%. In 1994, the wild Chinook in-river survival rate from Lower Granite Dam to McNary Dam was estimated at 47%, with virtually all removals occurring at McNary Dam since no operational return-to-river diversion system was present that year, so the fixed 50% expansion to LGR-equivalents on removals was proper in that year also. In post-1994 years, wild Chinook and wild steelhead had relatively small "raw" numbers of PIT-tag fish removed at downstream dams."

Wild Chinook	**1994**	**1995**	**1996**	**1997**	**1998**	**1999**	**2000**	**2001**	**2002**	**2003**
Pre-adj. C0 est. #[1]	3,621	2,725	1,919	682	3,081	4,469	6,573	233	6,410	9,001
Removal # Percent	910 25.1	8 0.29	1 0.05	1 0.15	0	0	41 0.62	1 0.43	60 0.94	60 0.67
Wild Steelhead				**1997**	**1998**	**1999**	**2000**	**2001**	**2002**	**2003**
Pre-adj. C0 est. #				454	776	1,113	1,871	103	4,107	3,343
Removal # Percent				0	13 1.68	0	0	0	9 0.22	12 0.36

Pre-adj. C0 est. # is the estimate prior to subtracting twice the removal number.

Page B-8 – BPA Comments:

- *Finally, an attempt is made at an explanation for the 50% survival probability used to deal with downstream removals. It is not sufficient, however. Why not use a dam-specific adjustment, rather than poking all downstream removals and assuming a common survival to every downstream dam? Have the effects of violations of this 50% survival assumption been examined? It is known that violations of this assumption occur, because survival between MCN, JDA, and BON is not 100%, so survival to one dam (e.g., at 50%) is not equivalent to survival to the other dams, as is implicitly assumed by using a single survival probability to all dams. Additionally, if using a single survival rates is warranted and if survival to BON is to be used each year, it should be possible to use the estimated survival to BON for the year, rather than assuming 50% survival each year.*

Response: Although dam-specific adjustments could have been used, the relatively low numbers of fish being affected as will be shown in response to the next BPA comment, makes all the concerns being raised here an over-reaction to a negligible effect.

- *Show the number of removal on a dam-specific basis that you contribute to d_0 and d_1.*

Response: The following two tables show the initial number of PIT-tagged smolts estimated in study categories C_0 and C_1 and final values obtained after the adjustment for fish removed at dams below LMN. In Table 1, the percent change from initial to final smolt estimate after the adjustment was minimal for wild Chinook, wild steelhead, and hatchery steelhead at less than 2%, except for wild steelhead in 1998 and wild Chinook in 1994. The high rate for wild Chinook in 1994 was due to no return-to-river capability that year at McNary Dam (all but two fish were MCN removals); the estimated reach survival from LGR to MCN was estimated at 47%, in line with the fixed 50% rate.

In Table 2, a higher percent change from initial to final smolt estimate after adjustment is seen for hatchery Chinook than was seen for wild Chinook or all steelhead. However, even these removal adjustment changes were generally less than 4%. The planned removals for physiological testing of PIT-tagged Chinook from Dworshak, Rapid River, and McCall hatcheries in the lower Columbia (mostly at Bonneville Dam) are the main reason for the higher percent change seen with these three hatcheries compared to Imnaha or Catherine Creek acclimation ponds. It should be noted that even if no survival rate expansion were applied, one would still, at a minimum, need to subtract the d_0 and d_1 fish removed below LMN in computing the final C_0 and C_1 smolt numbers. So relative to this minimum adjustment, the changes due to the CSS adjustments of $2*d_0$ and $2*d_1$ are effectively one -half the percentages shown in tables 1 and 2. The bottom line is that the CSS adjustment in years after 1994 has contained relatively small numbers of fish. Therefore, the suggestion of the reviewer that we should fine tune our adjustments to each dam where PIT-tag fish removals are taking place by using estimates of reach survival from LGR to that particular dam appears to be excessive. It would have relatively little effect on the resulting numbers of smolts estimated in C_0 and C_1 over the CSS approach.

Table 1. Change in C_0 and C_1 smolt estimates from initial to final value after adjusting for removals below LMN for wild Chinook and wild/hatchery steelhead.

Sp/RT Code[1]	Migr. year	Category C_0 smolt numbers				Category C_1 smolt numbers			
		final[2] C_0	initial C_0	remove d_0	change initial to final	final[2] C_1	initial C_1	remove d_1	change initial to final
WCH	1994	1,801	3,621	910	50.3%	4,431	8,459	2,014	47.6%
	1995	2,709	2,725	8	0.6%	14,206	14,260	27	0.4%
	1996	1,917	1,919	1	0.1%	5,209	5,213	2	0.1%
	1997	680	682	1	0.3%	1,936	1,936	0	0.0%
	1998	3,081	3,081	0	0.0%	12,276	12,296	10	0.2%
	1999	4,469	4,469	0	0.0%	26,140	26,150	5	0.0%
	2000	6,494	6,576	41	1.2%	16,833	17,051	109	1.3%
	2001	231	233	1	0.9%	20,307	20,589	141	1.4%
	2002	6,218	6,338	60	1.9%	12,687	12,911	112	1.7%
	2003	8,879	8,999	60	1.3%	12,694	12,846	76	1.2%
	2004	2,252	2,292	20	1.7%	16,504	16,698	97	1.2%
WST	1997	454	454	0	0.0%	2,984	2,990	3	0.2%
	1998	750	776	13	3.4%	5,150	5,374	112	4.2%
	1999	1,113	1,113	0	0.0%	6,992	6,992	0	0.0%
	2000	1,871	1,871	0	0.0%	10,616	10,616	0	0.0%
	2001	103	103	0	0.0%	11,892	11,932	20	0.3%
	2002	4,045	4,061	8	0.4%	8,726	8,802	38	0.9%
	2003	3,320	3,344	12	0.7%	7,132	7,160	14	0.4%
HST	1997	3,390	3,394	2	0.1%	19,095	19,113	9	0.1%
	1998	2,926	2,938	6	0.4%	17,958	17,998	20	0.2%
	1999	3,952	3,956	2	0.1%	20,975	20,983	4	0.0%
	2000	4,408	4,410	1	0.0%	18,804	18,808	2	0.0%
	2001	372	376	2	1.1%	19,132	19,226	47	0.5%
	2002	6,129	6,145	8	0.3%	14,038	14,110	36	0.5%
	2003	6,459	6,479	10	0.3%	10,118	10,144	13	0.3%

[1] Sp/RT is species and rear-type code: WCH = wild Chinook; WST = wild steelhead; and HST = hatchery steelhead.

[2] Final C_0 = initial C_0 - 2*d_0 and final C_1 = initial C_1 - 2*d_1.

Table 2. Change in C_0 and C_1 smolt estimates from initial to final after adjusting for removals below LMN for hatchery Chinook.

Sp/RT Code[1]	Migr. year	Category C_0 smolt numbers				Category C_1 smolt numbers			
		final[2] C_0	initial C_0	remove d_0	change initial to final	final[2] C_1	initial C_1	remove d_1	change initial to final
DWOR	1997	2,529	2,531	1	0.1%	3,613	3,613	0	0.0%
	1998	11,151	11,181	15	0.3%	13,128	13,214	43	0.7%
	1999	10,484	10,518	17	0.3%	19,083	19,207	62	0.6%
	2000	13,075	13,477	201	3.0%	5,416	5,580	82	2.9%
	2001	886	910	12	2.6%	16,872	17,480	304	3.5%
	2002	19,008	19,650	321	3.3%	14,914	15,570	328	4.2%
	2003	17,697	18,033	168	1.9%	6,715	6,985	135	3.9%
	2004	6,280	6,370	45	1.4%	14,009	14,195	93	1.3%
RAPH	1997	4,176	4,178	1	0.0%	6,843	6,845	1	0.0%
	1998	4,402	4,420	9	0.4%	13,597	13,691	47	0.7%
	1999	7,040	7,094	27	0.8%	14,456	14,602	73	1.0%
	2000	11,046	11,332	143	2.5%	5,248	5,406	79	2.9%
	2001	966	1,014	24	4.7%	15,989	16,631	321	3.9%
	2002	13,625	14,065	220	3.1%	14,854	15,436	291	3.8%
	2003	16,858	17,142	142	1.7%	7,055	7,195	70	1.9%
	2004	3,484	3,520	18	1.0%	12,776	12,928	76	1.2%
MCCA	1997	6,761	6,761	0	0.0%	9,272	9,274	1	0.0%
	1998	3,849	3,887	19	1.0%	12,816	12,886	35	0.5%
	1999	8,407	8,477	35	0.8%	11,391	11,527	68	1.2%
	2000	13,064	13,336	136	2.0%	4,485	4,565	40	1.8%
	2001	1,000	1,034	17	3.3%	15,536	16,040	252	3.1%
	2002	10,280	10,662	191	3.6%	12,315	12,787	236	3.7%
	2003	19,696	20,034	169	1.7%	8,669	8,817	74	1.7%
	2004	2,359	2,391	16	1.3%	16,297	16,489	96	1.2%
IMNA	1997	2,219	2,221	1	0.1%	3,785	3,785	0	0.0%
	1998	1,995	1,995	0	0.0%	6,335	6,335	0	0.0%
	1999	2,869	2,869	0	0.0%	5,084	5,084	0	0.0%
	2000	4,396	4,456	30	1.3%	2,254	2,286	16	1.4%
	2001	366	376	5	2.7%	6,939	7,043	52	1.5%
	2002	4,637	4,735	49	2.1%	5,135	5,253	59	2.2%
	2003	6,683	6,755	36	1.1%	2,908	2,936	14	1.0%
	2004	1,302	1,318	8	1.2%	4,456	4,502	23	1.0%
CATH	2001	379	391	6	3.1%	4,642	4,724	41	1.7%
	2002	2,445	2,499	27	2.2%	3,120	3,192	36	2.3%
	2003	3,201	3,247	23	1.4%	1,403	1,423	10	1.4%
	2004	503	513	5	1.9%	1,869	1,885	8	0.8%

[1] Hatchery Code is: DWOR = Dworshak; RAPH = Rapid River; MCCA = McCall; IMNA = Imnaha; and CATH = Catherine Creek.

[2] Final C_0 = initial C_0 - 2*d_0 and final C_1 = initial C_1 - 2*d_1.

- *"Estimation of SARs for study categories:" SAR1(T0) and SAR2(T0) have been discussed but not defined in the report. Define all measures.*

Response: $SAR_1(T_0)$ is a combination of dam-specific transport SARs in LGR-equivalents that are weighted by the proportion of run-at-large in total transportation occurring at each dam. $SAR_2(T_0)$ is the sum of returning adults from transported PIT-tagged fish divided by the sum of PIT-tagged smolts transported from each dam in LGR-equivalents. Parameter $SAR_2(T_0)$ is the primary SAR for evaluating transportation.

Page B-9 – BPA comments:

- *A "common annual routing rate to the raceways" was used -- what is this? What value was used?*

Response: A same rate of 2/3 PIT-tagged hatchery Chinook to raceways and 1/3 PIT-tagged hatchery Chinook returned to river for first-time detected fish from CSS participating hatcheries has been used at LGR, LGS, and LMN since 2000. It was accomplished by having the Separation-by-Code (SbyC) electronics at the Snake River collector dams divert 2 PIT-tagged fish to the raceways for every 3 PIT-tagged fish arriving from a particular CSS hatchery. In 2002 and 2003, the CSS coordinated with state and tribal researchers to divert ½ of their PIT-tagged wild Chinook to the raceways using SbyC. In 2004, this was increased to the same rate of 2/3 PIT-tagged wild Chinook and wild steelhead being routed to the raceways using SbyC. By utilizing a common annual routing rate for a group of PIT-tagged fish of interest, one achieves self-weighting across the three collector dams relative to their proportional contribution of each collector dam to total transportation. The benefit of achieving self-weighting is that $SAR_1(T_0)$ and $SAR_2(T_0)$ become equivalent in estimating the transportation SAR.

- *The notation used to define AC0 and AC1 is insufficient. It does not preclude using adults that were removed at downstream dams for any reason. Because many removed fish are not sacrificed, it is conceivable that some of these "removed" fish may return as adults. Are these adults included in AC0 and AC1?*

Response: The BPA review is mistaken. PIT-tagged smolts that are removed at downstream dams are considered permanently removed, regardless of whether sacrificed or not. For example, a fish detected only on the separator at McNary Dam later in the summer after the start of the transportation program of summer migrants would be considered as removed at that site, and therefore, any adult return from that particular fish would not be counted.

- *One assumes not, because this would positively bias the SARs for the C0 and C1 groups; however, the notation used implies that these removed fish are included in AC0 and AC1.*

Response: In the draft report we say "AC_0 = tally of adults of smolts that passed the three Snake River collector dams undetected (capture histories "1000AAAA" where A=0 signifies not being detected and A=1 signifies detection and return-to-river at a downstream site." If these fish had been removed at MCN, JDA, or BON, it would have been coded with a digit >1 in the site position of the capture-history table's field called CAPTURE_DI). Such a returning adult

would not have been tallied in AC_0. This same logic applies to "AC_1 = tally of adults of smolts that passed the three Snake River collector dams with at least one detection (capture histories "11AAAAAA" or "101AAAAA" or "1001AAAA" where the A=0 signifies not being detected and A=1 signifies detection and return-to river at a downstream site. If a returning adult has a CAPTURE_DI site-position digit where A>1 in the above capture-history list, then that adult will not be tallied in AC_1.

- *It looks like SAR2(T0) is used in this report for overall SAR of transported fish, rather than SAR1(T0), unless otherwise specified. Is this correct? Clarify.*

Response: Yes, $SAR_2(T_0)$ is the primary transportation SAR parameter. Table B-1 provides a summary of which annual reports utilized $SAR_1(T_0)$ (Annual Report 2001 for wild Chinook and 2002 for both wild and hatchery Chinook) and $SAR_2(T_0)$ (Annual Report 2000, 2001 for hatchery Chinook, 2003/04, 2005, and 2006) as the primary measure of transportation SAR. The clarification of why we returned in 2003 to using $SAR_2(T_0)$ as the primary transport SAR is detailed form the bottom of page B-8 through top of page B-9 in Appendix B of the 10-yr report.

Page B-11 – BPA comments:

- *The expected value of the size of the C1* group should be presented. At the least, the definition of the C1* group should be explained. It does not make intuitive sense to define it in terms of the T0, C1, and T0* statistics, because the T0 and T0* statistics are based on different groups of fish.*

Response: Contrary to what the reviewer suggests, the parameter T_0 and T_0* are based on the same underlying group of PIT-tagged fish. When this group if PIT-tagged fish are expanded to LGR equivalents and summed, we get the starting number of smolts in group T_0 at LGR. Further, expansion of this group allows us to estimate the number that would have been in T_0, which we call T_0*, provided T_0 fish had been transported at the same rate as the untagged run-at-large. In that situation, the population arriving LGR forebay of PIT-tagged fish of a particular CSS group, such as Rapid River Hatchery Chinook for example, would consist of C_0 fish "destined" to pass three collector dams undetected, T_0* fish "destined" to be collected and transported, and a remainder of fish that are "destined" to be collected and bypassed assigned to group C_1*. The sum of the T_0 and C_1 fish equal the collected portion of the PIT-tagged group. By subtracting the number of fish in T_0* from the sum of T_0 and C_1, we obtain an estimate of residual bypassed fish. In most years this is a very small, often immeasurable number, but in 1997 when the management action was to route many untagged fish, this group accounted for upwards of 25% of the run-at-large population of Chinook and steelhead.

B-12 – BPA Comments

- *The CSS states that "the rate of harvest is assumed independent of whether fish had been transported as smolts. [These] assumptions … apply to both TIR and D." Where does the CSS actually make use of this assumption? Is it only in their interpretation of results about TIR and D?*

Response: This assumption about harvest rate is utilized primarily when addressing losses during the adult fishes' upstream migration from Bonneville Dam to Lower Granite Dam. Although the rate of harvest is likely unaffected by whether smolts outmigrated in barges or inriver, the opportunity for harvest as transported fish may experience more straying effects could be another reason why we observed differential loss during the upstream migration based on prior downstream migration history. But even in the lower Columbia River prior to passing Bonneville Dam, there are opportunities for harvest in some years which we cannot directly measure with the PIT-tag data. Here again, we assume the rate of harvest is independent of prior downstream migration history. The effects of harvest removal will be to lower the magnitude of estimated SARs of both inriver and transported fish, but it will have less of an effect on those parameters that are based on the ratios of these two SARs (e.g., *TIR*s and *D*) if the harvest rates are independent of downstream migration history.

Reference:

Zar, J.H. 1984. Biostatistical analysis, 2nd Edition. Prentice-Hall, Englewood Cliffs NJ. 718 pp.