Responses to COMMENTS OF THE RESPONSIBLE PARTY:
Preassessment Data Report #10

Chronic Exposure of Seaducks to Oil Released by *the Selendang Ayu*
at Unalaska Island

Consistent with 15 CFR sec. 990.14 of the regulations governing natural resource damage assessments under the Oil Pollution Act, 33 USC sec. 2701, et seq., the Trustees provided the responsible parties with an opportunity to comment on this preassessment data report. The responsible parties' comments to that draft are below. **Trustee responses to the comments are in bold;** these responses are reflected in the final preassessment data report.

Page 2; Paragraph 2; Abstract

"In 2005 and 2006 the trend of P450 induction for harlequin ducks followed the trend in estimated degree of oiling from the Selendang Ayu, i.e. Skan Bay was the most heavily oiled and had ducks with highest induction rates, followed in order by Humpback Bay, Portage Bay and minimally oiled Chernofski Harbor."

> Insert at end of sentence ", however, within each year the difference in P450 values in ducks from Skan Bay, Humpback Bay, and Portage Bay were not significantly different from each other.
>
> This statement is not correct. The statistically insignificant "trend" of the means did not follow this pattern in 2006, where the order was Skan, Portage, Humpback then Chernofski.
>
> **#1. This study was not specifically designed to examine variation in levels of P450 inductions among oiled bays. If it were, sample sizes would have been higher within each oiled bay. These analyses and results were conducted and presented as general data descriptions. While it true that we concluded that there was no 'significant variation' among oiled bays, the power to detect such differences was poor given the small sample sizes. Given the confusion created by these analyses, we eliminated these results from this report.**

"Further, the second year results show that exposure continued more than a year after the spill."

> Replace with "Further, the second year results show that ducks in the spill area were exposed to EROD-inducing chemicals, possibly PAHs from the spilled oil, in the winter of 2005-2006.

#2. This sentence was modified for clarity along the lines suggested by the RP, but we did not include the qualifier that the source was "possibly from spilled oil". The source of exposure is a fixed conclusion as long as the assumptions of the study design are met.

"The observed levels of exposure are comparable to concentrations associated with reductions in survival following exposure from the Exxon Valdez spill."

This statement makes the presumption that there is a direct correlation between EROD activity in duck liver and the magnitude (concentration and duration) of exposure to PAH from the *Selendang Ayu* fuel and *Exxon Valdez* cargo and that the response to the two oils is the same; no data supporting this presumption has been presented in this report or any other Preassessment reports. Therefore, this statement should either be deleted or modified to accurately reflect the results of this Preassessment study.

Because of the manner of reporting liver EROD activity in this study and for similar studies in Prince William Sound, it is not possible to determine if liver EROD activity actually was similar in *Selendang Ayu* and *Exxon Valdez* studies.

Delete. The authors state that values are not comparable between years because of variability in EROD activity.

From page 5 of this PADR "…*we assumed that the absolute values of EROD activity were not directly comparable among years*.". It is logical that if comparison among years is not possible, comparison among years and locations also is not possible.

#3. All *direct* comparisons of P450 values with Exxon Valdez studies have been deleted.

"Although it is not known how long these effects will persist, studies following the Exxon Valdez suggest that they may persist for years."

This spill is different from the EVOS in type of oil, location, shoreline characteristics, oil behavior, and other factors. This is speculation and should be deleted from this PADR.

#4 This sentence references published literature and includes a qualifier (i.e., may) as such this is not speculation and this statement has been retained.

Page 3; Paragraph 3; Introduction

"*Extensive studies of harlequin ducks in Prince William Sound following the Exxon Valdez oil spill revealed that these ducks have high site fidelity and feed on intertidal resources making them susceptible to continuing oil exposure. The research has shown*

that chronic exposure to oil can have long-term deleterious effects on survival of harlequin ducks (Histrionicus histrionicus) (Esler et al. 2002)."

> Esler et al. (2002) reported that adult female harlequin ducks from the path of the *Exxon Valdez* oil spill (EVOS), had lower winter survival than females from an unoiled area of western Montague Island. Harlequin ducks from the heavily oiled northern Knight Island area had higher liver and blood EROD activity than ducks from Montague Island. Induced (elevated) EROD activity is a biomarker of exposure to one or more of several classes of organic chemicals, including the polycyclic aromatic hydrocarbons (PAH) in petroleum. They attributed the poor winter survival of female ducks to a reduction in fitness of the ducks due to exposure to EVOS PAH.
>
> Keith Parker (personal communication), a statistician with the Data Analysis Group, observed that differences in the over-winter survival of female harlequin ducks at oiled and unoiled areas actually were nominal, and not statistically significant. Esler and Jarvis (2000) report over-winter survival and standard errors for adult females at oiled and unoiled areas of 78.0% (± 3.3%) and 83.7% (± 2.9%), respectively. A 95% confidence interval is approximately two standard errors, and for oiled and unoiled areas, 95% confidence intervals are (71.4, 84.6) and (77.9, 89.5). Confidence limits overlap and, therefore, there is no statistically significant difference in survival at $\alpha = 0.05$. A test for equal survival at oiled and unoiled areas yields an approximate p-value = 0.20 (for a two tailed test, which the Exxon Valdez Trustees typically use). Esler and Jarvis (2000) did not report results for testing the hypothesis of equal survival at oiled and unoiled areas.
>
> The language used to describe these results on survival became more definitive over time. Esler and Jarvis (2000) "speculate" and "suggest that oil exposure" is the cause of differences in winter survival. Based on these and 2001 results, Esler et al. (2002) find these data "strongly support" lower female survival in oiled areas. By 2003, Peterson et al. (2003) found "significant implications for population trajectories" from these same estimates of survival. Peterson et al. (2003) report results in terms of mortality, not survival. Presented as mortality estimates, differences look more dramatic, 22% for oiled vs. 16% for unoiled. However, the more likely scenario—the one supported by the data—is that such nominal differences were due to sampling error and spatial variation.
>
> Integral (2006) reported that: "The winter survival study was repeated in 2000–2003 (Bodkin et al. 2003). Preliminary findings of this unpublished study indicate that there are no significant differences in cumulative winter survival of female harlequin ducks between previously oiled (81 percent survival) and unoiled (84 percent survival) areas. Female survival was not reevaluated in 2005 when apparent exposure levels were higher (based on EROD activity) in oiled areas than in unoiled areas."

Rosenberg et al. (2005) also performed a study for the EVOS Trustee Council on population dynamics of harlequin ducks in spill path and non-spill path areas of PWS between 1997 and 2005 and concluded: "Annually, we observed a slight increase in recruitment. We found no significant difference in the change in density (trends) between oiled and unoiled treatments (P = 0.761) and the mean rate of change for oiled areas (0.0125, P = 0.138) and unoiled areas (0.0186, P = 0.304) was not significantly different from zero. The lower proportions of females in oiled areas provided the only evidence for a possible lingering oil spill effect. Demographic data interpreted in concert with other biological parameters leads us to conclude that harlequin duck populations are recovering from the *Exxon Valdez* oil spill."

Thus, the modeled winter survival for female harlequin ducks from oiled areas of PWS (Esler et al., 2002) probably was not different from winter survival in unoiled reference areas.

#5. This is a lengthy comment which while obviously focused on several sentences in this report, really is most relevant to other published literature (i.e., Esler et al 2002). There are two main flaws in this comment. First, the comment misapprehends the *Information Theoretic* approach to data analyses. The foundation of this approach is fundamentally different from hypothesis testing approaches using frequency based statistics. It is not appropriate to mix the 2 approaches as was apparently done by Keith Parker. Esler and Jarvis (2000) never said that there was a significant difference in survival of Harlequin Ducks in oiled and unoiled areas (in fact there never tested this *hypothesis*). What they did was assess the support for models in which survival was different in oiled and unoiled areas as well as models where there was no difference in survival between oiled and unoiled areas. They found greater support for models where survival differed between areas and report the estimates as such. Following-up such AIC based model selection with hypothesis testing approaches is not valid.

The comment regarding the language used to describe these results (in combination with additional data) refers to a paper published in *Science*. The appropriate scientific venue for such comments would be in a rebuttal paper published in that journal. Such a rebuttal would be subjected to peer review (as was the original paper), which is the accepted measure of scientific rigor.

The final sentence of this comment is based on an incorrect statistical interpretation of published data and results.

Page 3; Paragraph 4; Introduction

"*To determine whether seaducks were likely experiencing ongoing injury as a result of the Selendang spill, we assessed petroleum hydrocarbon exposure through a comparative analysis of cytochrome P450 induction in liver tissue.*"

Cytochrome P450 analysis is a biomarker for exposure to inducing chemicals, including PAHs. As such, it is not a measure of whether harlequin ducks "were likely experiencing ongoing injury."

As this is stated to be a Preassessment Data Report, we suggest the following changes:

"We used Cytochrome P450 levels in duck liver as a biomarker to determine if seaducks were being exposed to PAH in the *Selendang Ayu* spill area in the two years after the spill."

#6. Based on published data, there is a relationship between Cytochrome P450 induction and injury. As such, this statement is a logical description of the overall point of this study. Statement was modified along the lines suggested by the RP. The connection between exposure and injury is now only addressed in the Discussion.

Page 3; Paragraph 6; Methods

"*After capture, birds were transported a short distance to a mobile field laboratory where they were banded and examined. Their weight, sex, and age (adult or immature) were recorded*."

Information on weight, sex, and age of the captured ducks were recorded. This information did not seem to be used although there is some research to indicate that these variables need to be taken into account in terms of establishing potential negative effects from sub-lethal exposure to residual oil. Further, EROD activity in Harlequin ducks has been shown to have very high natural variability, which again may be attributed to covariates such as age, diet, and sexual stage of development for these birds (Lee and Anderson 2005).

#7. True that these data were not utilized in these analyses. However, the publication by Lee and Anderson focuses on fish and has no data relative to this issue for birds. We lack sufficient sample size to control for all of these additional variables in our analyses. The key issue in this study design is that we are assuming that all else except for the Selendang Ayu oil exposure is equal between the oiled bays and Chernofski. Thus if we were capturing significantly different samples with regard to age and sex cohorts, AND there are differences in P450 induction across these variables, then the results might be spurious relative to the effects of oiling. To date, there are no data indicating variation in P450 inductions for Harlequin Ducks by age and sex.

Further, there were no differences in sex ratios, or mean body mass at capture among oiled and unoiled bays, there was a slight difference in age ratios in 2006. Thus, data would support the assumption that samples obtained are equivalent. Text regarding this assumption has been added to the report.

Page 4; Paragraph 3; Analyses

"*The P450 data from wild birds were highly skewed with a few very large values, so we log transformed those data prior to analyses to improve normality*."

The authors note that the P450 data were highly skewed with a few very large values so they log transformed the data prior to the analyses to improve normality. We are not sure why they made the transformation since their analyses are based on randomization tests. This class of tests allows inference without requiring assumptions about underlying distributions. At the same time, if the underlying distribution is important, a plot of the transformed data and/or results of a goodness fit test for normality should be included.

#8. It is true that randomization tests do not require the assumption that the data are normally distributed. That is, the tests are valid when data are normally distributed and when they are skewed. The issue here is how does one both analyze such skewed data and logically present the results. When data are highly skewed, interpretation of the mean is nebulous and the relationship between the standard error and the 95%confidence interval for the mean is no longer valid. That is, the mean plus 1.96 times the standard error no longer represents the upper 95% confidence limit for the mean. We decided, *a priori*, to analyze and present reasonably normally distributed data such that there would be direct comparability between the results and the data summaries presented.

Page 5; Paragraph 3; Results

"Cytochrome P450: The measured enzyme (or P450) response of dosed mallard eggs varied among years. In 2005 the ratio of EROD activity in dosed to control eggs was 62. In 2006, the ratio was only 24.2. Therefore, we assumed that the absolute values of EROD activity were not directly comparable among years. For presentation in figures, we used the ratio of the response difference across years to adjust P450 values."

The use of a constant correction factor ignores variability by treating the correction factor as a constant, when in fact it is a random variable. As a general comment, the procedures utilized through out this document appeared to discount variability, thus, potentially making "differences" appear more substantive than they actually are.

EROD activity was extremely variable in birds from oiled and reference bays and between 2005 and 2006. Because of the inter-annual variability in the laboratory protocol for measuring EROD activity in oiled and reference duck eggs, it was not possible to make inter-annual comparisons. The investigators resorted to reporting results as the ratio of the average oiled site EROD activity to the reference site activity (assuming that any ratio greater than 1 was an indication of exposure to oil) within a year and adjusting values from different years based on EROD activity in oiled and unoiled duck eggs.

The presentation of these data in Figures 3 and 4 is confusing and makes it difficult to compare results with those from other studies, such as this in Prince William Sound (which Flint et al., 2007 do repeatedly). What are the units for "Log P450 induction" on the y axis? It is essential to present real values ((pmole/min/mg protein) so readers have some appreciation of the magnitude of induction in oiled and reference animals.

It is interesting to note that the investigators, in an earlier study that included EROD activity in harlequin ducks from the eastern Aleutians in the winters of 2002-3 (Miles et al., 2007), did not appear to have as much trouble with extreme variability in EROD activity in ducks from contaminated and clean locations. EROD activity was reported in their Figure 2 as pmol/min/mg protein. It would have been informative to compare the results from Miles et al. (2007) to those from the *Selendang Ayu* and *Exxon Valdez* spill sites. The EROD assays in the Miles et al. and Flint et al were performed in the same laboratory (U. Cal. Davis), presumably making results (and laboratory problems) similar.

EROD activity was used in studies after the *Exxon Valdez* oil spill to document continuing exposure of harlequin ducks to petroleum PAH (Trust et al., 2000; Esler et al., 2002). As in the *Selendang Ayu* study, ability to interpret trends between years and among study sites was hampered by extreme variability in values for EROD activity. Esler (2005), who managed the EVOS oil exposure and effects studies with harlequin ducks, explained in his 2005 GEM proposal to the EVOS Trustee Council: "Measurements of cytochrome P4501A (P450) have proven to be extremely useful for quantifying the degree of exposure to hydrocarbons following the EVOS for a number of vertebrates, including harlequin ducks. However, the ability to document interannual changes in exposure for harlequin ducks is eroded by dramatic differences in average P450 values between years, both for oiled and unoiled areas. There is no reasonable biological explanation for these widely differing values among years and we speculate that these are the result of differences within the laboratory. Because the P450 data are so critical for documenting changes in oil exposure over time, as well as for linking individual survival with oil exposure, we propose to concurrently reanalyze all archived HADU samples."

Esler (2005) continued: "Interpretation of P450 data for harlequin ducks has been hampered by dramatic interannual differences in EROD activity results. As

described in Fig. 1, average EROD activity reported for oiled areas ranged from 40.2 to 1981.8 pmol/min/mg across years and, for unoiled areas, the range was from 36.0 to 1187.9 pmol/min/mg across years. These almost certainly do not reflect real differences in exposure over time. Not only is the magnitude of differences among years in oiled areas far beyond what one would expect, one should predict that average exposure in oiled areas would decline over time with diminishing availability of oil, as has been described for other species. Also, there is no biological explanation for the dramatic differences across years in unoiled areas. One would expect that average EROD activity should remain stable over time in the unoiled areas. Also, interannual differences are fairly consistent between areas when considered across years; e.g., for both areas results are more than 30 times higher in 2001 than 2002. We are left to conclude that dramatic interannual differences are the result of variation in the laboratory processing."

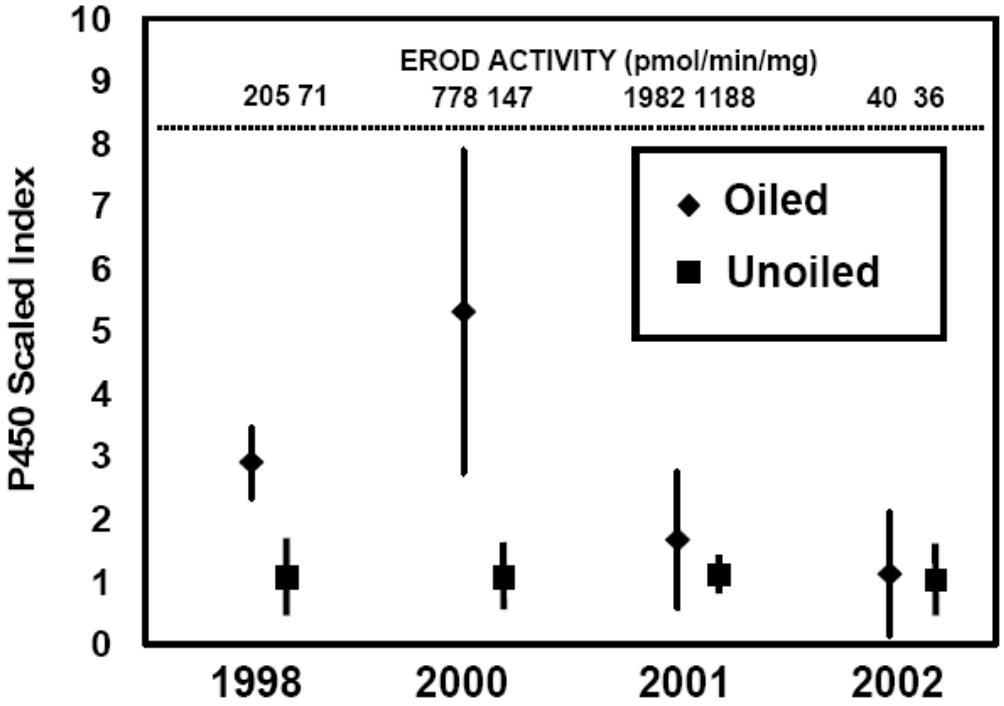

Figure 1. Cytochrome P450 (as measured by EROD activity) in harlequin ducks from oiled and unoiled areas of Prince William Sound, Alaska. The values above the dotted line are the average EROD activity reported from the lab. The figure scales the EROD data, setting the unoiled area at 1 for each year, to illustrate the change in the ratio of oiled:unoiled over time. From Elser (2005).

We have plotted the un-normalized EROD data from Esler (2005) to give a better idea of the tremendous interannual variation in EROD activity in harlequin ducks (Figure 2). There are no temporal trends when the data are presented in this way. EROD activity in ducks from oiled and unoiled locations is higher in 2001 than at

any other time. Activity is the same in birds from oiled and unoiled locations in 2002, implying a lack of exposure to inducing chemicals, possibly EVOS PAH.

Esler (2005) reported that: "Results from studies of captive harlequin ducks at the Alaska SeaLife Center corroborate the hypothesis of lab-induced interannual differences. During 2 winters (2000 and 2001) female harlequin ducks were captured from an unoiled area and held from September to March in captivity. In each winter, ducks ingested oil in controlled amounts and their P450 response was measured at season's end. Despite similar, controlled handling and dosing of ducks, as well as sample handling, between years, dramatically different results were reported in the 2 years. Mean EROD activity of birds was 634.6 and 2239.4 pmol/min/mg in 2000 and 2001, respectively. More surprisingly, EROD activity of control birds was 86.7 and 235.3 pmol/min/mg in 2000 and 2001, respectively. The ratio of EROD activity for oiled:control birds was similar between years (7.3:1 and 9.5:1 in 2000 and 2001, respectively), suggesting that the magnitude of the differences was valid, but that values could not be directly compared across years."

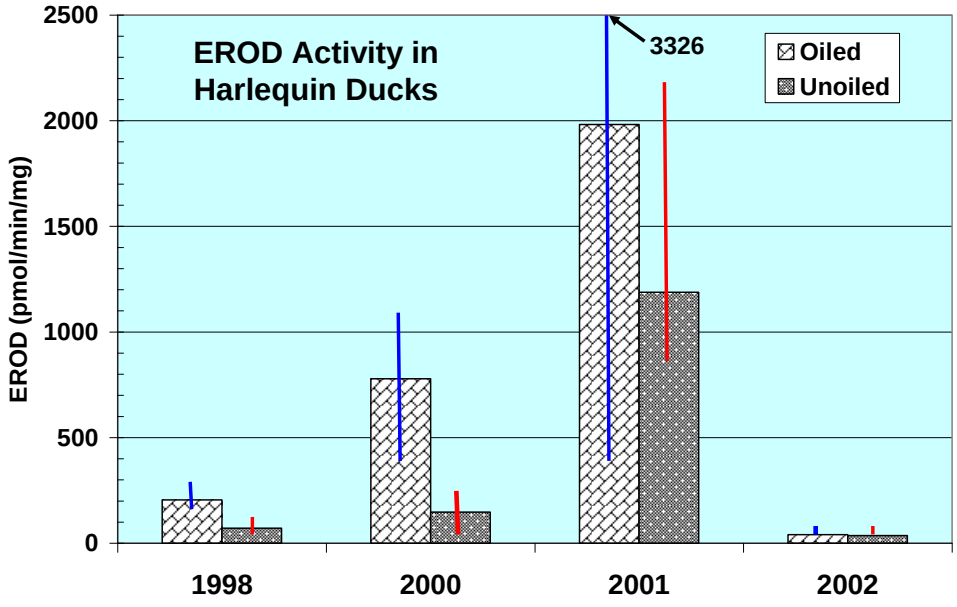

Figure 2. Mean EROD activity in tissue of harlequin ducks from oiled and unoiled locations in Prince William Sound between 1998 and 2002. This presentation shows the large inter-annual variation in EROD activity in ducks from oiled and unoiled locations.

#9. While true, this treating of the correction factor as a constant only applies to the presentation for the figures. It has absolutely no effect on the statistical tests and associated results. We could graphically present these data any number of ways. Real values can easily be approximated from the

figures by simple exponentiation of the values. Given the remainder of the comments on this topic, the use of a correction factor at all becomes moot. In this final report version, we now simplify the overall presentation based on the study design where an effect persists until there is no longer a difference between P450 induction values between oiled and reference (i.e., Chernofski) areas. Accordingly, any further consideration of comparison of absolute or log P450 values and associated correction factors no longer applies.

Page 5; Paragraph 5; Results

"*There was no difference in levels of relative EROD activity among oiled bays (i.e., Skan, Portage and Humpback) in either year (2005: P = 0.294, 2006: P = 0.716). When oiled locations were pooled, there was a significant difference in EROD activity between oiled and unoiled areas in both years (2005: P = 0.001, 2006: P < 0.002) (Figure 3). Between 2005 and 2006, the relative level of EROD activity in harlequin ducks sampled in the same bays did not change. Ducks continued to show levels of hydrocarbon exposure consistent with the degree of oiling in the four bays. Further, the relative decline over from one year to the next was minimal (Figure 4).*"

As mentioned earlier, a constant is being used in place of a random variable. Such calculations ignore the variability in the random quantity and further ignore propagation of the variability through final error bar determinations.

The harlequin duck exposure studies based on liver EROD activity in harlequin duck liver were technically flawed. Flint et al. (2007) did not document any deleterious sublethal effects to harlequin ducks from the *Selendang Ayu* spill.

Pooled EROD activity in birds from the 3 more heavily oiled bays was claimed to be significantly higher than in birds from the lightly oiled bay. A simple fixed effects ANOVA for the 2005 data using log(P450) as the response and location as the factor, shows the only significant difference is between Chernofski and Skan. This is a departure from treating the locations as random effects, which by their selection based on levels of oiling were not random effects. These data draw in question the value of the pooled data comparison.

There was no change in mean EROD activity from 2005 to 2006 in birds from any oiled or reference site. One would expect that concentrations of bioavailable *Selendang Ayu* oil PAH in sediments or prey would decline in the year between samplings, with a corresponding decline in EROD activity in birds from oiled sites if the *Selendang Ayu* oil actually was the cause of induction.

While there may be a visible "trend' in the P450 induction (mean value) between Skan, Humpback and Portage Bays, the results are apparently not statistically significant based on the author's own analysis. Given the variability among the data values, the results are statistically indistinguishable and other speculative

comments assuming significance should be stricken. Ducks did not show EROD production consistent with oiling in 2006 where the order was Skan, Portage, Humpback and then Chernofski.

#10. It is true that the results from the 3 oiled bays were not statistically significant, but we are randomizing the sample among the 3 oiled bays then dividing each by a constant (i.e., Chernofski) for each specific year. Thus, the division by the Chernofski value allows inclusion of both years data in a single test. Dividing all 3 oiled bays by a constant should have no effect on the ultimate test statistic as both the real data and the randomized data were treated consistently. Further, if the net effect of this division by a constant is as described by the reviewer, then this underestimates the variation among bays. Yet even with this potential reduced variation, we found no difference among bays. Thus, increasing the variation will not change our conclusion that exposure rates did not vary among oiled bays. So the comment had no effect on the statistical result. Nonetheless, all comparisons among oiled bays have been eliminated as they were not part of the original study design.

Conversely, we believe this comment is technically flawed. Beyond that it is true that we did not document any deleterious sub-lethal effects, only elevated P450 values, we reference published literature relative to the deleterious effects.

I think the reviewer misinterprets how a randomization test would be conducted on a 'pooled sample'. If we are lumping non-significantly different populations (2 areas) with a single population that was different, this would reduce our ability to detect a difference between oiled and unoiled areas. What we conducted would be the functional equivalent of an overall ANOVA looking at the effect of oiling. It appears that the reviewer is looking directly at the results for pair-wise comparisons.

Contrary to the reviewer's assertions, the report does not contend that there was no change in mean EROD activity from 2005 to 2006, as this would require use of a constant correction factor (the reviewer's interpretation may be based on his interpretation of the graphics). Rather, we stated that the <u>pattern</u> observed was consistent among years. We tested for annual variation in the pattern of variation in P450 values and failed to reject the null, this is the result we used to make the statement that there was no difference in the pattern of P450 exposure. Nonetheless, given confusion regarding these tests, we eliminated all comparisons among oiled bays from this report.

<u>Page 6; Paragraph 3; Discussion</u>

"Studies in Prince William Sound following the Exxon Valdez spill indicate that levels of exposure similar to those observed here correlate with reduced over-winter survival of harlequin duck females (Esler et al. 2002). Thus, mortality estimates based solely on dead birds found immediately after the spill event likely underestimate the magnitude of the overall mortality as birds chronically exposed to low levels of hydrocarbons also suffer increased mortality."

EROD activity in harlequin ducks sampled after the *Selendang Ayu* spill from oiled and reference bays on Unalaska Island apparently was much lower than in harlequin ducks from oiled and reference areas of PWS, measured by Esler (2005). Because of the extreme variability in EROD activity in ducks from oiled and reference areas in both studies, plus the inappropriate normalization used, it is difficult to assess whether the differences in EROD activities in ducks from PWS and from Unalaska Island really were different. If differences in EROD activity are real and not a laboratory artifact and if EROD activity actually is a biomarker of exposure to oil in both cases, exposure of ducks in the spill zone of the *Selendang Ayu* spill is much less than exposure of ducks in the spill zone of the *Exxon Valdez* spill. The purported effects of lingering exposure to EVOS PAH in PWS, a decrease in winter survival of females (Esler et al., 2002), are equivocal, as discussed above; therefore, the effects of the apparently much lower exposure of harlequin ducks to *Selendang Ayu* PAH are unlikely.

A confounding factor in assessing long-term exposure to and effects of the *Selendang Ayu* oil spill in harlequin ducks is that the coastal waters were and still are not free from chronic petroleum contamination. It is estimated that more than 3,000 ships per year pass near Unalaska Island along routes similar to that of the *Selendang Ayu* (pers. comm., R. Morris, Capt. USCG). Releases of oil in bilge or ballast discharges are not unusual in the area. In addition, Miles et al. (2007) show that there are duck liver EROD-inducing chemicals in Dutch Harbor, apparently not associated with oil spills.

It should be noted that background oiling of birds and beaches in the Aleutian Islands has been documented (Byrd et al 1995)[1]. While the study focused on the western and central Aleutian Islands, the following excerpts are relevant to this PADR:

> "Effects of beach oil on wildlife was not measured, but we found oil on live and dead birds of nearly every species that occurs in the nearshore zone."
>
> "Nearly 1,000 km away in the eastern Aleutian Islands, observers found 11 dead oiled birds at Yunaska on one sample beach between 28 May and 4 June."

[1] Byrd, G.V., J.C. Williams and G. Thomson. 1995. The status of oil pollution on beaches of the Alaska Maritime National Wildlife Refuge, 1992-1994. Alaska Maritime National Wildlife Refuge, Homer, AK.

"The fact that we found no oil spots during brief surveys on narrow transects at particular islands should not be interpreted as evidence that no oil was present"

"The samples that we had analyzed indicated the majority was probably bunker C, a fuel used in many commercial ships today and the main fuel used for ships during WWII."

The findings of Byrd et al. (1995) was confirmed in the study area by the documentation of oil not consistent with the *Selendang Ayu* source oil (Figure 3) in the study area and elsewhere.

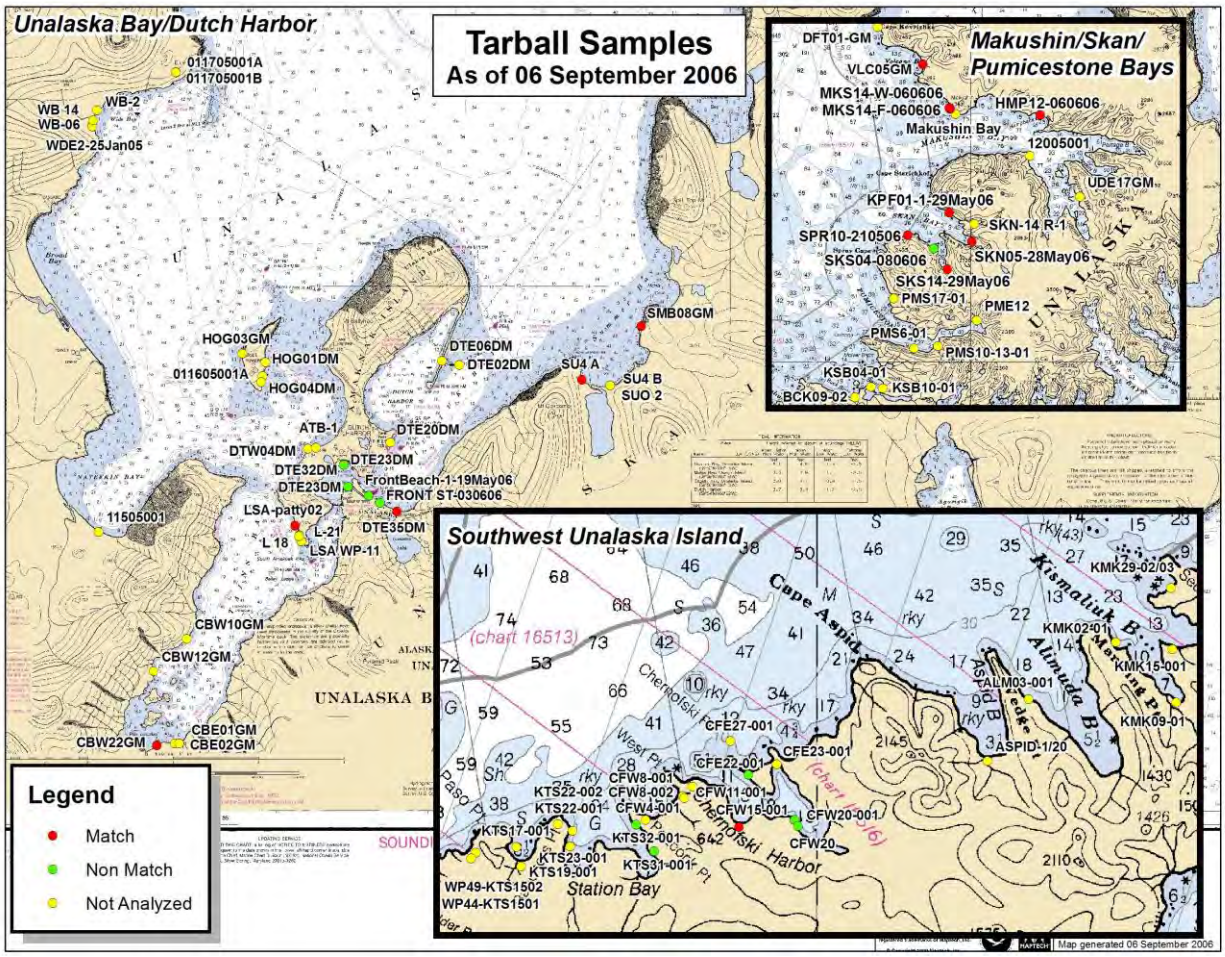

Figure 3. Results of source allocation of Tarball Samples in the vicinity of the *Selendang Ayu*.

#11. We eliminated the portion of this paragraph that claimed similar exposure rates as PWS. The remainder of the paragraph was retained in the final report, as it validly cites to published literature. Contrary to the reviewer's contentions, log transformation would not invalidate our results.

Again its true that the randomization tests does not require normality, but the test is just as valid on normalized data as skewed.

The accepted scientific standard would be for the RP to publish either data contrary to those of Esler 2002, or some sort of rebuttal/commentary. Thus, the conclusion advanced by the RP that these results are equivocal based on a non-peer reviewed assessment, is not scientifically valid.

While true that this area may have been exposed to non-Selendang Ayu oil, this cannot explain the difference in P450 induction between oiled and unoiled areas. The RP's data on tar-ball allocation shows non-Selendang tar balls in Chernofski harbor. Our analyses assumes that non-Selendang oil, if present; is uniformly distributed between oiled and reference areas. It would appear that the RP's data support this assumption.

Page 6; Paragraph 4; Discussion

"*More than a year following the wreck of the Selendang Ayu, levels of P450 induction exhibited a pattern that followed the degree of oiling among bays documented during the SCAT process*."

Within each year, the pattern is not statistically significant among oiled bays. This statement should be stricken or corrected.

From page 5 of this PADR "…*we assumed that the absolute values of EROD activity were not directly comparable among years*.". It is logical that if comparison among years is not possible, comparison among years and locations is not possible.

#12. The test for the pattern of variation is still valid even when we do not assume that the absolute values are comparable among years. See description of the randomization test. The fact that the RP disagrees with statement that the overall pattern of P450 induction seems to relate to pattern of oiled from the SCAT surveys represents an opinion. Unfortunately, the SCAT data on oiling cannot be quantified and thus, this relationship cannot be examined statistically. However, the pattern was consistent among years and it is appropriate in the discussion of data to present all potential explanations for observed patterns in these data. However, given that the SCAT assessments cannot be quantified, we eliminated all text relating P450 exposure to SCAT assessments from this report.

"*Given the known effects of oil exposure on survival of harlequin ducks (Esler et al. 2002), we hypothesize that harlequin ducks were still dying as a result of exposure to hydrocarbons released by the Selendang Ayu more than a year after the spill. Although it*

is not known how long harlequin ducks in the spill area will be exposed to Selendang oil, based on studies conducted in response to other spills, and depending on factors such as the geomorphology of beaches and oil toxicity, the spill effects may persist for years."

These statements are hypothetical and speculative, as such, it is not appropriate for inclusion in a Preassessment Data Report.

The biology of harlequin ducks probably is quite different in PWS and around Unalaska Island, making such extrapolations highly problematic. Much less is known about the biology of harlequin ducks around Unalaska Island than in PWS. Females may be under lower energetic stress in the winter in the southern Bering Sea than in PWS. Reed and Flint (2007) reported that harlequin ducks wintering near Dutch Harbor required a relatively low foraging effort during periods of high energetic demand (winter), suggesting availability of abundant, high-quality food. Esler et al. (2002) observed that harlequin ducks wintering in western PWS must feed nearly continuously to meet metabolic needs. There is no documentation that harlequins in the *Selendang Ayu* spill area were stressed or experienced increased mortality as a result of the spill. Thus, it is much less likely that female harlequin ducks still are stressed and still experience lower winter survival in oiled than unoiled areas of Unalaska Island 4 years after the *Selendang Ayu* oil spill than was purported for PWS, 10 or more years after the *Exxon Valdez* spill.

#13. In the absence of direct data from this spill, we relied on effects from published literature dealing with this same species to reach a logical conclusion. While true that there is no documentation that Harlequins in the spill area experienced increased mortality, there is similarly no documentation that they did not. In fact, published data suggests that birds with elevated P450 values do suffer higher mortality. If you have any data to the contrary, or citations to published data relevant to this issue, please provide. The remainder of this comment appears to be based on speculation.

Page 10; Figure 3

"*P450 induction was significantly higher in oiled locations than in Chernofski Bay ("reference") (P < 0.002) in both 2005 and 2006. Data from 2006 were adjusted prior to plotting to account for among year differences in EROD activity of dosed duck embryo standards (i.e., 62/24.2 = 2.54).* "

This is another case of ignoring variability by treating the correction factor as a constant, when in fact, it is a random variable.

#14. Data were only adjusted for purposes of plotting. Such adjustment had no effect on statistical results. Such adjustment has now been removed from the presentation.

Page 10; Figure 4

"*Data from 2006 were adjusted prior to plotting to account for among year differences in EROD activity of dosed duck embryo standards (i.e., 62/24.2 = 2.54)*."

Again, a highly variable random quantity is being treated as a constant. The correction factor of 2.54 is a ratio estimator derived from two other ratio estimators that have their own inherent errors. In 2005 the sample sizes were small, there were 4 inductions and 2 controls and in 2006, there were 5 inductions and 4 controls, again each with their own inherent variability. This is another case of a treatment of a constant without consideration of the propagation of variability.

#15. Same as #14.

Responsible Party's References

Bodkin, J., B. Ballachey, D. Esler, and T. Dean. 2003. Patterns and processes of population change in selected nearshore predators. EVOS Project No. 423. Prepared for the *Exxon Valdez* Oil Spill Trustee Council. U.S. Geological Survey Alaska Science Center and Center for Wildlife Ecology, Simon Fraser University.

Byrd, G.V. and L. Daniel. 2007. Preassessment Data Report: Bird Species Found Oiled, December 2004 – January 2005, at Unalaska Island Following the M/V *Selendang Ayu* Oil Spill. US Fish and Wildlife Service, Alaska Maritime National Wildlife Refuge, Homer, AK.

Esler, D. 2005. Quantifying temporal variation in harlequin duck exposure to Exxon Valdez oil. GEM Proposal, EVOS Trustee Council, Anchorage, AK.

Esler, D., T. D. Bowman, K. A. Trust, B. E. Ballachey, T. A. Dean, S. C. Jewett, and C. E. O'Clair. 2002. Harlequin Duck population recovery following the "Exxon Valdez" oil spill: progress, process, and constraints. Mar. Ecol. Prog. Ser. 241: 271–286.

Esler D and Jarvis RL. 2000. Winter survival of adult female Harlequin Ducks in relation to history of contamination by the *Exxon Valdez* oil spill. J. Wildl. Manage. 64(3):839-847.

Flint, P.L., J.L. Schamber, K.A. Trust, A.K. Miles, and B.W. Wilson. 2007. Preassessment Data Report: Chronic exposure of seaducks to oil released by the M/V *Selendang Ayu* at Unalaska Island. U.S. Geological Survey-Alaska Science Center, Anchorage, AK. 16 pp.

Integral Consulting, Inc. 2006. Information Synthesis and Recovery Recommendations for Resources and Services Injured by the *Exxon Valdez* Oil Spill. *Exxon Valdez* Oil Spill

Restoration Project 060783 Final Report. Report to *Exxon Valdez* Oil Spill Trustee Council, Anchorage, AK. October 2006. 377 pp.

Lee, Richard F. and Jack W. Anderson. 2005. Significance of cytochrome P450 system responses and levels of bile fluorescent aromatic compounds in marine wildlife following old spills. Marine Pollution Bulletin. 50:705-723.

Miles, A.K., P.L. Flint, K.A. Trust, M.A. Ricca, S.E. Spring, D.E. Arrieta, T. Hollmen, and B.W. Wilson. 2007). Polycyclic aromatic hydrocarbon exposure in Steller's eiders and harlequin ducks in the eastern Aleutian Islands, Alaska. Environ. Toxicol. Chem. (in press).

Peterson, C.H., S.D. Rice, J.W. Short, D. Esler, J.L. Bodkin, B.E. Ballachey, D.B.Irons 2003. Long-term ecosystem response to the Exxon Valdez Oil Spill. Science, 302:2082-2086.

Reed, J.A. and P.L. Flint. 2007. Movements and foraging effort of Steller's eiders and harlequin ducks wintering near Dutch Harbor, Alaska. J. Field Ornithol. 78:124-132.

Rosenberg, D.H., M.J. Petrula, D.D. Hill, and A.M. Christ. 2005. Harlequin Duck Population Dynamics: Measuring Recovery from the *Exxon Valdez* Oil Spill. *Exxon Valdez* Oil Spill Restoration Project 040407 (through 050407) Final Report. Report from Alaska Dept. of Fish and Game, Division of Wildlife Conservation, Anchorage, AK. 54 pp.

Trust, K. A., D. Esler, B. R. Woodin, and J. J. Stegeman. 2000. Cytochrome P450 1A induction in sea ducks inhabiting nearshore areas of Prince William Sound, Alaska. Marine Pollution Bulletin 40:397-403.