N00204.AR.002433
NAS PENSACOLA
5090.3a

COMMENTS FROM U S NAVY IN RESPONSE TO FLORIDA DEPARTMENT OF ENVIRONMENTAL PROTECTION COMMENTS TO DRAFT REMEDIAL INVESTIGATION REPORT ADDENDUM NUMBER 2 SITE 2 NAS PENSACOLA FL
2/22/2002
U S NAVY

Navy Response to Florida DEP Comments on the Draft Remedial Investigation Report Addendum Site 2 Waterfront Sediments, NAS Pensacola
April 19, 2002

Comments received 22 February 20 02

Comment 1:
Page 2-3: A sample location map should be added to the document.

Response 1:
An orientation map is provided in Figure 1-1, which identifies the study area and the two reference locations. Figure 2-1 identifies the decision units at Site 2.

Comment 2:
Page 3-11, 1st paragraph: Define or explain the Hazard Quotient for this table.

Response 2:
Agreed. The definition for Hazard Quotient will be included as a footnote to Table 3-1 and will also be explained in the text.

Comment 3:
Page 3-11, Table 3-1: Define "Sediment Benchmark" in the footnotes.

Response 3:
Agreed. The definition for Sediment Benchmark will be included as a footnote to Table 3-1.

Comment 4:
Page 3-14, Section 3.1.2 Sediment Subsurface: The subsurface findings explained in Appendix B need to be discussed further. Appendix B discusses subsurface contamination for metals, Polynuclear Aromatic Hydrocarbons (PAHs) and Semi Volatiles. Figures showing subsurface contamination with a top view and a vertical cross section should be included in this section.

Response 4:
Agreed. Subsurface findings will be addressed in the body of the report.

Comment 5:
Page 3-15, 1st paragraph: What is this sentence trying to say?

Response 5:
Section 2.1.3 explains acid volatile sulfides and simultaneously extracted metals (AVS/SEM) and the role they play in the equilibrium partitioning approach. AVS/SEM provides general information on the possible bioavailablity of five (Cd, Cu, Ni, Pb, and Zn) distinct metals. Section 3.1.3 provides the results from the AVS/SEM analyses. The last sentence in that section states: "For Site 2 the AVS/SEM results revealed that metal constituents are generally not bioavailable, though they may be at Stations GH-12, CD-23, and GH-67." We will insert the word 'present' into the statement so

it reads, "...though they may be present at Stations GH-12, CD-23, and GH-67". Table 4-3 shows no impact on biota at Stations GH-12, CD-23, and GH-67.

Comment 6:
Page 4-15, 1st paragraph: Are we ignoring evidence and data that would lead to a Feasibility Study?

Response 6:
The Navy assumes this comment refers to paragraph 4.4.10; Station CD-23. This station was one of the furthest sampling stations from Site 2, and was the closest station to the wet slip on the west side of the Allegheny Pier. Boat traffic in and out of the wet slip may have contributed or might be the source of the contamination at Station CD-23. Table 3-1 shows that 14 constituents with HQs above 1 were detected at Station CD-23. However, 11 of these exceedances were for PAHs and metals, indicating the contamination is likely related to vessel traffic in the Site 2 area.

Comment 7:
Page 5-1: Does the decision tree really take us in the direction of NFA?

Response 7:
Table 4-6, Page 4-10, shows that 9 of the 11 decision units within Site 2 follow the steps to NFA when applied to Flow Chart 2, the Simplified Decision Flow for Each Decision Unit. Decision Units CD-23 and EF-45 are the exceptions. The multiple lines of evidence gathered during the investigation of Site 2 indicate that the area is recovering from past Naval Base activities.

Comment 8:
Appendix B, Page 2: A sample location map for the sediment and subsurface samples is required for a complete review of this material.

Response 8:
Appendix B was compiled by USEPA Region IV and cannot be revised. An orientation map is provided in Section 1, and Figure 2-1 includes station decision unit identifications. In addition, Appendix A Figure 1 shows the sampling grid and sediment sampling plan. For further clarification, Step 7 of the DQO process (Page 21) provides a verbal explanation of the sampling plan, procedures, and rationale.

Comment 9:
Appendix B, Page 7: Metals contamination found in the subsurface for DUs 3 and 4 should be discussed in Chapter 3 of the RI Addendum.

Response 9:
Agreed. Subsurface sediment will be discussed in the RI Addendum.

Comment 10:
Appendix B, Page 9, paragraph 2: This paragraph should be discussed in chapter 3 of the RI Addendum.

Response 10:
Agreed. Subsurface sediment will be discussed in the RI Addendum.

Comment 11:
Appendix B, Page 12: PAHs contamination found in the subsurface for DUs 3 and 11, should be discussed in Chapter 3 of the RI Addendum.

Response 11:
Agreed. Subsurface sediment will be discussed in the RI Addendum.

Comment 12:
Appendix B, 3.1.4 Pesticides/PCB Analysis: Why is there no discussion of the subsurface findings in this section?

Response 12:
The Draft RI Addendum focused on the upper 6 inches of sediment. Subsurface sediment will be discussed in the RI Addendum. Appendix B was compiled by USEPA Region IV under contract to the Navy and cannot be revised.

Navy Response to USEPA Comments on Draft Remedial Investigation Report Addendum Site 2 Waterfront Sediments, NAS Pensacola April 19, 2002

Comments received 7 January 20 02

Comment 1:
Page 1-2, fourth Paragraph - The information contained in this report validates the conclusions of the previous RI to conduct a feasibility study to provide the information for the proper selection of the appropriate remedial action for this site.

Response 1:
The purpose of this study was to discern if conditions adverse to benthic communities in Pensacola Bay Site 2 area exist today (2000). The original RI Report was completed June 1997. Following the RI, an FS was completed and finalized in September 1997. During the public comment period (8 Dec 97 to 22 Jan 98), a comment was received questioning the preferred alternative. The comment suggested that the Navy should do something or do nothing at the site. Since the regional area had been directly impacted by three hurricanes after the original study, the Partnering Team agreed to reevaluate the current conditions. The multiple lines of evidence gathered during the investigation of Site 2 indicate that the area is recovering from past Naval Base activities.

Comment 2:
Page 2-2, Figure 2-1 Map showing the 150' X 150' Sampling Grids for this investigation - Identify the sample grids by the code used in the text (e.g., EF-45).

Response 2:
Agreed. Figure 2-1 will be updated to include station decision unit identifications.

Comment 3:
Page 3-1, 3.1.1 Field Chemistry Results - The comparison of maximum values from the previous study to the results of the composite samples in the present study may overemphasize the differences in the contaminant concentrations when the composite sample values (which may represent an average value for an area) is lower than the previous maximum value. This comparison may also under emphasize the differences in the contaminant concentrations when the composite sample value (which may represent an average value for an area) is lower than the previous maximum value. This comparison may also under emphasize the differences in the contaminant concentrations when the composite sample value is greater than the previous maximum value.

Response 3:
Agreed. Comparisons between the two sampling events have been deleted. Instead, 2000 sediment data will be compared to the reference concentration (developed for the study), and the lower of sediment screening values (USEPA Region IV or FDEP) for sediment benchmark comparisons. Table 3-1 shows the surface sediment chemistry and Table 3-2 shows the subsurface chemistry,

each of which compare the constituents to the reference location and the sediment benchmark values.

Comment 4:
Page 3-22, 3.2.5 Benthic Community Results, First Paragraph - The mixture of pollution tolerant and pollution sensitive organisms may be due to the sample collection occurring during a recruitment period (samples were collected in March) of the benthic community seasonal cycle.

Response 4:
The Tier 1 members and the eco-sub group labored for over one year finalizing the details of the study plan, and received buy-in from all members (Tier 1, eco-group, and even a Tier 2 representative). The sampling schedule determined by USEPA was conducted at their earliest convenience (March).

The purpose of the diversity sampling was not to conduct an in-depth benthic community assessment as that would require seasonal, if not monthly, sampling over a series of years. The intent of the Pensacola Site 2 sampling was to provide a snap shot of the organisms currently making up the benthos. In addition, the information collected is included in the weight of evidence approach which utilizes the sediment quality triad approach (species diversity, toxicity tests, and chemical analyses) to address the original question, 'Are conditions adverse to benthic communities in Pensacola Bay Site 2 area in existence today (2000)?'

Benthic community from the Site 2 area will be compared to the reference stations (18 and 22) as agreed upon in the DQO document.

Comment 5:
Page 4-3, 4.2.1 Interpretation of Biological Endpoints to the Triad - The "non-normalized" toxicity test information should be incorporated into a redrafted decision making triad table.

Response 5:
Agreed. Non-normalized data will be presented throughout the report.

Comment 6:
Page 4-12, 4.4.3 Station GH-12 - How can the SEM/AVS ratio be above 1 but the SEM-AVS value be below 0.0?

Response 6:
Table 4-6 shows the SEM-AVS for Station GH-12 is 0.095. This will be corrected in 4.4.3, the narrative portion of this decision unit, in the document.

Comment 7:
Page 5-1, 5.0 Conclusions and Recommendations - Sub-lethal effects are appropriate endpoints for remedial ecological risk assessments. The evidence appears to contradict the second paragraph.

Response 7:
The Navy will address uncertainties identified by reviewers of the document to better support earlier conclusions. The sublethal endpoints are included in the evaluation of site conditions and have been evaluated using the criteria established in the approved DQO document (Appendix A).

Navy Response to NOAA Comments on Draft Remedial Investigation Report Addendum Site 2 Waterfront Sediments, NAS Pensacola
April 19, 2002

Comments received 23 January 2002

Comment 1:
The report's conclusions and recommendations (§5.0) are not technically supported. Section 5.0 of the report concludes that, based on comparisons to 1996 data, conditions at Site 2 "continue to improve" as evidenced by: 1) decreasing surface sediment chemistry constituents, 2) no acute and minor sub-lethal effects in the sediment bioassays, 3) "vastly improved" benthic community, 4) SEM:AVS data. Based on these conclusions, the Navy recommends No Further Action at Site 2.

These conclusions and the ensuing recommendation do not appear to be supported by results in the report. Moreover, substantial uncertainties, not discussed in the report, greatly decrease the confidence associated with a No Further Action recommendation at this time. Technical results and important uncertainties are discussed below.

Response 1:
Comparisons to previous sampling events have been deleted from the text. The 2000 sampling effort was designed to answer the question, Are conditions adverse to benthic communities in Pensacola Bay Site 2 area in existence today (2000)? Therefore, the conclusion will be based on the current data set. However, the results from the 1996 event have been summarized in Section 1.

Comment 2:
Sediment Chemistry

The conclusion that sediment chemistry at Site 2 "continue to improve" is not apparent from the results presented. Rather, the data suggest levels of sediment contamination in 2001 were more or less the same as that observed in 1996. This observation of minimal change between 1996 and 2001 is not unexpected for the following reasons: 1) the Industrial Waste Water Treatment facility was installed over 25 years ago, 2) major contaminant inputs presumably ceased at that time and 3) the system has likely reached some level of equilibrium.

Response 2:
The purpose of this study was to discern if conditions adverse to benthic communities in Pensacola Bay, Site 2 area, exist today (2000). Therefore, comparisons to previous events have been deleted.

Comment 2a:
Table 3-1 "Chemistry Comparison 1993-2000" (§3.1.1), appears to be the primary basis for the report's conclusion that site sediment chemistry has diminished. However, this table has limited

value for temporal comparisons and may possibly be misleading for two reasons. One, in most instances, there are no data available for temporal comparisons; note the preponderance of "No Data" entries. Two, when data are available, the *highest* historical value was used. This comparison biases results towards a conclusion of *decreasing* sediment concentrations. (What station is represented by the data at the top of page 3-5, Table 3-1?)

Response 2a:
Agreed. The comparisons to previous events have been deleted.

Comment 2a (continued):
Consider taking a broader, site-wide contaminant-specific approach for evaluating temporal trends in surface sediment chemistry. For example, one could compare all 1996 cadmium surface sediment concentrations to all 2001 surface sediment cadmium results. Then one could statistically test the null hypothesis that there is no significance difference in cadmium concentrations between 1996 and 2001. There appear to be sufficient data in both years to provide a rigorous statistical comparison. Repeat this comparison for all site COPCs including perhaps ER-Ms. One would then have a technically sound basis for concluding whether surface sediment chemistry at Site 2 is increasing, decreasing or staying about the same.

Response 2a (continued):
The 1996 data was used to focus the effort and investigation for the 2000 Remedial Investigation. The application of the DQO Process enabled a comprehensive study design which would allow conclusions to be drawn from composited sediment data, toxicity tests, and species diversity determinations, instead of comparisons to previously collected data. The Navy acknowledges the potential uncertainly surrounding comparisons between collection techniques from 1996 versus 2000. To remove uncertainty, the 2000 Remedial Investigation focuses on the data collected in 2000.

Comment 2b:
The subsurface sediment chemistry data in Appendix B are not presented nor discussed in the main body of the report. Not determining the vertical nature and extent is a major uncertainty in this (or any) RI. The lack of vertical nature and extent represents a very large data gap for risk managers who must consider the feasibility of remedial alternative.

To fill this data gap, consider taking the broader, site-wide contaminant-specific approach described above for evaluating temporal trends surface sediment chemistry. However, instead of comparing 1996 and 2001, compare surface (0-6 inches) and subsurface (>6 inches) data. Differences in core lengths as well as speculations regarding compressed core samples can be discussed in the uncertainty section. Without this delineation, one will not have any insight regarding potential self-capping processes, sediment mixing or if deeper material is more contaminated than surface sediment. These findings would also significantly affect the feasibility of remedial alternatives.

Response 2b:
The Navy concluded conditions at Site 2 supported a diverse benthic community and therefore did

not focus on subsurface sediments. The Draft Remedial Investigation Report Addendum Site 2 Waterfront Sediments focussed on the condition of surface sediments at Site 2 and if those conditions are acceptable based on sediment chemistry, toxicity, and species diversity. Additional information about chemical nature and extent was collected to support any feasibility study that might follow this investigation. Subsurface sediment data will be included in the body of the report. Sub surface collection efforts were not completed at each of the desired locations due to compaction, this should be noted as there will be data gaps. A comparison of surface and subsurface contaminant concentrations is provided in Section 4.

Comment 2c:
Table 4-4 accurately describes the four categories of presumptive toxicity based on sediment guidelines developed by NOAA and FLDEP (Correct citation is missing in the report.). However, the Site 2 sediments appear to have been misclassified. For example, no comparisons to FLDEP guidelines are made. All Site 2 sediments have one or more chemicals which exceed FL TELs. For that reason alone, no sediments can be placed in Category I as the subject report has done. Table 4-5 shows ERM exceedences as well as Total and Mean ERM quotient. Similar tables should be constructed for FLDEP TELs and PELs as well as NOAA ER-Ls before Site 2 sediments can be properly categorized.

Response 2c:
Surface and subsurface chemical constituents were compared to the ERLs and ERMs, and quotients developed to determine into which category the decision unit should be classified (Section 4). In addition, the Navy applied surface (Table 3-1) and subsurface (Table 3-2) chemical constituents to the Florida PELs and TELs (Appendix F). These data were then used to classify DUs for their input to the sediment quality triad according to their chemical characteristics relative to sediment guidelines. To conservatively estimate potential excess risk, the Navy considered Categories 2, 3 and 4 to be "+" in the triad matrix, even though there is considerable uncertainty regarding toxicity for Categories 2 and 3.Toxicity test results further validate classification of many of the sediments as Category 1 at Site 2.

Comment 2d:
Sediment organic carbon was to be analysed in all samples. However, the data were "lost in the analysis process" (§3.1.4). A value of 0.1% TOC is assumed based on the 1996 data. If accurate, sediment-associated contaminants must be assumed to be highly available for uptake by biota. Consequently, the NOAA and FLDEP guidelines (which are generally based on silty sediments) become *less* conservative. Is a value of 0.1% reasonable given the low sand/high fine combinations shown in Table 3-3? Can we evaluate grain size results between 1996 and 2001 to compare differences/similarities and thus, the reasonableness of the 0.1% TOC assumption?

Response 2d:
The SAP outlined that the sediment samples would also be analyzed for total organic carbon. Instead, the samples were analyzed for organic and non-organic % for material greater than 2 mm

in size and less than 2 mm in size. Data was salvaged by conservatively estimating 50% of the organic material less than 2 mm in size for each DU and is presented in Table 3-5.

Comment 2e:
The Total PAH expressions in Table 3.8 of §3.1.3.2 in Appendix B should be brought forward into the main body of the text. These results may help explain some of the observed toxicity. Is there a land-based source for these elevated PAH concentrations?

Response 2e:
Total PAHs for surface and subsurface sediments will be presented and discussed in the body of the report. The source for these PAHs are primarily from high water craft (military, commercial, and recreational) traffic in site area, and the influence from stormwater.

Comment 3a:
Sediment Toxicity
Survival results in both the amphipod and mysid toxicity tests have been "normalized" for variations in the multiple control treatments (Tables 3-5 and 3-6). The "normalized" results are then carried forward into the sediment quality triad analysis. This "normalization" adjustment should be omitted for several reasons. One, this data manipulation was not specified in the Work Plan where statistical analysis of the toxicity data is described. Two, adjustments for control survival is typically not conducted for simple sediment bioassays. Rather, the control treatment is used as a performance standard to judge whether the test itself is valid. Normalization may be appropriate for very large programs such as EMAP but not for simple hypothesis testing like Site 2. Three, "normalization" was not consistently applied. Survival data were "normalized" but growth and reproduction results were not. As a result, the triad analysis contains both "normalized" and non-normalized results ("apples and oranges"). Four, the normalization procedure diminishes the clarity of the report. "Normalized" survival is greater than 100% in a number of the sediment treatments. These anomalous results will likely elicit questions from readers.

Response 3a:
Non-normalized data will be presented throughout the report.

Comment 3b:
No true sediment control was used in the Site 2 sediment bioassays. The control treatment for sediment bioassays consists of substrate the test organisms were either cultured in or collected in. The Site 2 bioassays used neither. Rather, an uncharacterised field-collected sediment was used to "dilute" sediments for the mysid bioassay and serve as the sole basis for all statistical comparison in both sediment bioassays. Almost nothing is known about the control sediment. No chemistry or grain size information is reported. Even the collection date and location are unclear. At a minimum, the location for collecting these sediments should be shown on a map. The lack of information regarding the sediment used as the control treatment is a potentially significant source of uncertainty.

Response 3b:
A true sediment control was used in all the sediment Site 2 exposures. Control sediment characterization (physical and chemical) will be included in the report, as will the GPS coordinates and a map identifying the sampling location.

Comment 3c:
Multiple control treatments were used in both the amphipod and sediment bioassays. As long as the variance among controls is small, the corresponding uncertainty associated with multiple controls is also small. This uncertainty should be evaluated in the subject report. In addition, control results for the both the amphipod and mysid bioassay tables are not reported (Tables 3-5 and Table 3-6, respectively). Results from the multiple control treatments should be reported in these tables as they are the sole basis for statistical comparisons.

Response 3c:
Control data for each batch of tests conducted will be included and an explanation of negative controls will be added to inform the reader.

Comment 3d:
Growth data for amphipods and mysids were not statistically analysed if survival was significantly affected. The triad analysis requires input from both the lethal and sublethal responses. Therefore, all sublethal data should be statistically analysed regardless of survival.

Response 3d:
Agreed. Sublethal data will be statistically analyzed and included in the body of the report.

Comment 3e:
Mysid reproduction was measured during the laboratory bioassay (see contractor report in Appendix C) but was not reported in the main body of the report (see mysid summary Table 3-6). Table 3-6 should report results of the statistical analysis of both reproduction endpoints; i.e., percent females and percent females with eggs. A statistical analysis of all mysid reproduction endpoints is required by the triad analysis (see top of page 4-2, §4.1.1 and Table 4-2).

Response 3e:
Agreed. Mysid reproduction data will be statistically analyzed and included in the body of the report.

Comment 3f:
Feeding regime for the two sediment bioassays requires clarification. The Work Plan indicates test organisms were to be fed (see Tables 1 and 2 in Appendix A). The contractor's reports in Appendix C are silent with regards to feeding. EnSafe has verbally indicated mysids, but not amphipods,

were fed. The contractor's report should be revised to indicate what feeding regime was actually used in both bioassays.

Response 3f:
Agreed. The report will be modified to indicate that mysids were fed during the 7-day exposure, and amphipods were not fed during their 10-day exposure. An addendum will be added by the contractor, to report the feeding regime used.

Comment 3g:
Considerable effort was devoted to evaluating the toxicity of "diluted"sediments in the mysid bioassay (Appendix C). What was the value of these additional test treatments and how are the results used? The DQO portion of the Work Plan indicates these data would be used to determine NOECs and LOECs (Figure 1, Appendix A).

Response 3g:
Eco group members discussed the importance of collecting as much data as possible so informed decisions could be made. A modification was made to the mysid chronic test method, in which sediments were tested at 100, 50, and 25%. These concentrations were made by volume-to-volume dilutions with the control sediment. The purpose of these concentrations was to develop gradient data, if effects were noted in the higher concentrations. Since no effects were observed in survival of the 100% in the mysid tests, the development of NOECs and LOECs was not necessary and was therefore not included in the body of the report.

Comment 3h:
Control chart with the positive reference toxicant suggests amphipods used to evaluate the toxicity of Site 2 sediments were the most resistant ever tested by this contract laboratory (see contractor's report in Appendix C). This finding, and its potential impact on characterizing Site 2 sediment toxicity, should be discussed in the report.

Response 3h:
While the sensitivity of the *Leptocheirus* reference toxicant test was on the high end, LC_{50} values were well within the upper and lower 2 standard deviations. As stated in the DQO document, 'reference toxicant tests indicate the relative sensitivity of the test organisms being used and demonstrate a laboratory's ability to obtain consistent test results with the test method' as was confirmed by the control chart in the contractor's report found in Appendix C. Therefore, there was no impact to the sediment tests.

Comment 4:
Benthic Community Analysis
Changes observed in the benthic community between 1996 and 2001 do not appear to warrant a conclusion of "vastly improved". The small differences observed may simply reflect normal

seasonal variability. The 1996 samples were collected in winter (January 1996) while the 2001 samples were collected during the spring biological recruitment period (March 2001). This alternative alone could explain the observed differences and should be discussed in the report. Also, the potential influence of collection methodologies on differences observed in 1996 and 2001 should be noted and discussed.

Response 4:
The Tier 1 members and the eco-sub group labored for over one year finalizing the details of the study plan, and received buy-in from all members (Tier 1, eco-group, and even a Tier 2 representative). The sampling schedule determined by USEPA, and was conducted at their earliest convenience (March 2000).

The purpose of the diversity sampling was not to conduct an in-depth benthic community assessment as that would require seasonal, if not monthly, sampling over a series of years. The intent of the Pensacola Site 2 sampling was to provide a snap shot of the organisms currently making up the benthos. In addition, the information collected is included in the weight of evidence approach which utilizes the sediment quality triad approach (species diversity, toxicity tests, and chemical analyses) to address the original question, 'Are conditions adverse to benthic communities in Pensacola Bay Site 2 area in existence today (2000)?'

Benthic community from the Site 2 area will be compared to the reference stations (18 and 22) as agreed upon in the DQO document.

Comment 5:
SEM:AVS
The molar ratio of simultaneously extracted metals to acid volatile sulfides (SEM:AVS) can be used as a theoretical measure of bioavailability for five divalent metals (Cd, Cu, Ni, Pb, Zn). Theory predicts that at ratios less than 1, AVS in pore water will bind all 5 divalent metals sufficiently to reduce or eliminate biouptake and thus toxicity. At ratios greater than 1, insufficient AVS is present to bind all metal present in pore water and toxicity may be expressed.

The SEM:AVS model was developed largely with silty (high TOC) sediments. Because the report has no TOC results and has assumed a very low TOC (0.1%), some discussion should be devoted to appropriateness of using the SEM:AVS model in low TOC sediments.

The subject report is misleading when it concludes, "bioavailability of metals in the sediments is limited based on AVS SEM data" (page 5-1). An equally plausible alternative interpretation is that when SEM:AVS ratios exceed 1 (as they do in three locations, Table 3-2), amphipod growth or survival is adversely affected (Table 3-5).

Response 5:
AVS/SEM analysis is a tool which provides information on possible bioavailability of the five distinct metals listed above. The AVS normalization approach assumes that select trace metals (cadmium,

copper, nickel, lead, and zinc) bind to sediment sulfide, thus becoming insoluble. Those metals not bound to the sediment sulfides may be bioavailable to organisms. If the concentration of SEM is less than the molar concentration of AVS, all metals should precipitate as metal sulfides and not be bioavailable. Conversely, if SEM exceeds AVS, then free metal ions may exist in pore water. The analysis assumes that when the difference between SEM and AVS (SEM-AVS) concentration is greater than 1 or the ratio of SEM/AVS is greater than 1, bioavailability from the select divalent cationic metals is probable.

Because the Site 2 sediments were relatively low in TOC content, AVS/SEM analyses may be less important for defining site conditions than other aspects of this investigation. AVS/SEM information is presented only to enhance the study.

Comment 6a:
A good base map should be prepared for this report. No figure currently exists that shows the relationship between the 1996 grid locations and the 2001 Decision Unit locations. Such a map would clarify temporal trend analysis. It would also greatly facilitate communication of results (e.g., where was toxicity observed). The base map must show the location of the submerged seawall. Once plotted, the seawall may emerge as an important feature for interpreting RI results and managing risks at Site 2.

Response 6a:
Agreed. Maps indicating 1996 stations, 2000 stations, and the submerged sea wall will be added to this document. Comparisons between 1996 and 2000 data have been deleted from the report Conclusions for Site 2 will be based on the current data set.

Comment 6b:
Be consistent in station location nomenclature. Text, figures and tables all vary between grid designations (e.g., CD-23) and Decision Unit designations (e.g., DU 11) when describing station locations. This greatly diminishes the clarity and transparency of the document. This reviewer suggests grid designations are illuminating, especially when displayed on a good base map.

Response 6b:
Agreed. Consistency noted and will be applied to the entire report.

Comment 6c:
In Figure 1-1, indicate which location is reference station 18 and which is station 22.

Response 6c:
Agreed. Stations 18 and 22 will be identified in Figure 1-1.

Comment 6d:
In Tables 3-4 and 3-7, indicate the meaning of the range of values shown under "*Notes:*". Clarify what is meant by "* = Stations with species diversity overlap" in Table 3-7. Does it mean that stations with asterisks are different from other stations, or not different from each other, or what?

Response 6d:
Agreed. The footnotes will be clarified. With regard to species diversity overlap (asterisks), an explanation will be provided in the report. The eco-group may remember during the sampling design discussions, costs associated with individual diversity analyses were kept in the fore front and to reduce costs the group agreed to 'share' diversity data when corners of the decision unit matrix overlapped one another. Stations which 'shared' data due to this overlap are denoted by an asterisks. This aspect will be best explained with a small diagram.

Comment 6e:
Report a variance term (SD or SE) for each mean toxicity value in Tables 3-5 and 3-6. Also, report means and variances for the multiple controls.

Response 6e:
Agreed. Standard deviations and standard errors will be presented for each mean toxicity value in Tables 3-5 and 3-6. Means and variances will be reported for multiple controls.

Navy Response to University of Florida Comments on Draft Remedial Investigation Report Addendum Site 2 Waterfront Sediments, NAS Pensacola

We have reviewed at your request the Remedial Investigation Report Addendum, Site 2 Waterfront Sediments, Naval Air Station Pensacola, Florida prepared by Ensafe, Inc. The document presents results of a study designed to determine whether sediment conditions have improved since the last investigation performed in 1996. Previous investigations in 1993 and 1996 found Hazard Indices (HI) greater than 10 at several locations within a general impacted area encompassing approximately 5.4 acres. These HIs represented potential toxicity of the whole sediment and were calculated using as benchmarks the Florida Threshold Effects Levels (TELs) of the Florida Sediment Quality Assessment Guidelines, or the USEPA Sediment Screening values when the former were not available. The hazards were driven by metals and bis(2-ethylhexyl)pthalate (BEHP), which were historically released into Pensacola Bay from a plating shop housed in building 71.

The most recent field investigation was conducted in 2000 and used the sediment triad approach to assess sediment quality within eleven 150' x 150' Decision Units (DU) located in the general impacted area, and on two nearby reference DUs. Samples for toxicity testing and chemical analysis were derived from a composite of 6-10 samples from each DU. Samples for benthic community analysis were three discrete samples collected at the center and near two corners of each DU. In addition, a 36" sediment core was collected from eight of the 11 DUs to establish if contamination extended beyond 6"; the depth encompassed by all other samples.

Concentrations of contaminants in these samples were compared with the maximum concentration recorded at each DU during the 1993 and 1996 sampling events. In addition, the data were used to calculate a condition score, which was based on the results of sediment chemistry, toxicity tests, and benthic assessment for each DU. The report concludes that conditions at Site 2 continue to improve and recommends no further action.

We have the following comments on this RI addendum:

Comment 1:
The report presents in Table 3-1 chemical data collected in 1993, 1996, and 2000, and the respective Hazard Quotients. The accompanying text repeatedly indicates that the data suggest a decreasing trend in contaminant levels from 1993 to 2000. However, this comparison may be flawed in that the 1993 and 1996 are the maximum value measured from discrete samples collected at each DU, while the 2000 data are from composites of 6-8 individual samples and are thus more representative of the average concentration. The apparent reductions in concentrations are small enough that they could be explained simply by a comparison of maximum versus average comparisons for these areas, rather than an actual loss of contaminant. Unless the data are compared on a uniform basis, it will be difficult to reach conclusions about trends in concentration over time.

Response 1:
Comparisons between the 1996 and 2000 data have been deleted. Instead, a history of the site and the 1996 findings will be discussed in Section 1, and the 2000 study will be discussed in depth in Sections 2, 3, 4, and 5.

Comment 2:
An inspection of the chemical data shows that seven of the 11 DUs still have HI values above 40. Again, it must be kept in mind that these HI were calculated using a composite of samples collected at eight locations evenly spaced throughout the approximately 0.5-acre plots. Concentrations at some areas within the plot could have HI values substantially above 40.

Response 2:
Because chemical constituents impact organisms differently, hazard indices (HI) have not been calculated for the data. Instead, mean ERM quotients have been used to assess the potential for toxicity.

Comment 3:
The study also determined the ratio and difference between Simultaneously Extracted Metals (SEM) and Acid Volatile Sulfides (AVS). This approach is based on the observation that AVS can bind metals in anaerobic sediments, thus reducing their bioavailability. The study found that all but three stations (viz., GH-12, GH-67, and CD-23) had SEM/AVS ratios <1, implying that divalent metals may not be bioavailable. We caution that this approach is only meaningful for anaerobic sediments; and aeration of sediments (such as during dredging operations) can reduce the pH causing release of heavy metals (Christensen, E.P., Metals, acid-volatile sulfides, organics, and particle distributions of contaminated sediments. Water Science & Technology, 37:149-156, 1998). We suggest that a discussion of this shortcoming of the method be included in the uncertainty section of the report.

Response 3:
AVS/SEM analysis is a tool which provides information on possible bioavailability of the five distinct metals listed above. The AVS normalization approach assumes that select trace metals (cadmium, copper, nickel, lead, and zinc) bind to sediment sulfide, thus becoming insoluble. Those metals not bound to the sediment sulfides may be bioavailable to organisms. If the concentration of SEM is less than the molar concentration of AVS, all metals should precipitate as metal sulfides and not be bioavailable. Conversely, if SEM exceeds AVS, then free metal ions may exist in pore water. The analysis assumes that when the difference between SEM and AVS (SEM-AVS) concentration is greater than 1 or the ratio of SEM/AVS is greater than 1, bioavailability from the select divalent cationic metals is probable.

Because the Site 2 sediments were relatively low in TOC content, AVS/SEM analyses may be less important for defining site conditions than other aspects of this investigation. AVS/SEM information is presented only to enhance the study.

Comment 4:
Chemical data were used in the sediment quality triad differently than in previous investigations. Instead of calculating hazards with respect to the Florida TELs, the report instead uses the Effects Range Medium (ERM) of the NOAA sediment quality guidelines as the comparison point (see Table

4-5). These data are then entered into the triad by classifying stations into four categories, depending on the degree and number of exceedances to the ERM and PEL values.

We agree with the final categories assigned to the stations during the inclusion of chemical data into the sediment quality triad (see pg. 4-8), except for station KL-34. Although BEHP is one of the most significant site-related chemicals of potential concern, it was not included in this comparison and in the resulting triad, probably because there is no ERM available for this chemical. If the PEL value were used, station KL-34 would have also been assigned to category 2 based on a calculated HQ for BEHP of 4.15. (Note: BEHP concentrations at stations IJ-12 and CD-23 also exceeded the PEL.). We recommend using the BEHP PEL for this portion of the triad and reclassifying station KL-34.

Response 4:
Comparisons to PELs and TELs is provided in Appendix F. In addition, a comparison of the mean ERM quotient results and the mean PEL quotient results is provided in Section 4. None stations had concordance between the two methods. Three stations (GH-12, IJ-34 and KL-34) showed an increase from Category 1 to Category 2 using the mean PEL quotient method. However, both toxicity tests showed no lethal or sublethal effects when compared to controls for stations IJ-34 and KL-34. GH-12 showed a sublethal effect for growth in the Leptocheirus test, but the overall input into the triad would remain the same (—) based on the scoring regime established in the DQO document for toxicity. Of note, EF-23 would be assigned a Category 3 because of the number of PELs exceeded (7). However, the input into the final triad matrix for sediment chemistry would not be changed .

Comment 5:
The report includes results of two sediment toxicity tests: 10 day Leptocheirus plumulosus survival and growth bioassay, and 7-day Mysidopsis bahia survival, growth, and reproduction test. The 10-day test identified station EF-45 as having a significant, although minor, effect on survival. Stations IJ-12, GH-12, GH-67, and both reference locations had a significant effect on growth. On the other hand, the 7-day M. bahia test did not identify any station as being toxic. Biological analyses also included benthic community studies that calculated three commonly used indices of biological diversity; Shannon Weiner Diversity, Pielou's Eveness, and Margalef's Species Richness indices. None of these parameters indicated adverse effects on benthic diversity at any of the stations. The design and execution of these studies appears to be reasonable, and suggests that sediment contamination is not having significant effects on populations found at the site.

Response 5:
The Navy agrees.

Comment 6:
The results, as well as anecdotal data included in the report, suggest that sediments have moved to the west. Station KL-34 is located in an area not previously sampled and had the highest

concentration of BEHP. It is not known if the impacted area extends beyond the boundaries of this station, especially to the north and south.

Response 6:
The study area focused on the locations with the highest hazard indices as identified in the 1996 investigation. The area of Station KL-34 was sampled during the Phase IIA event of the 1996 sampling. The sediment chemistry results did not indicate the potential for excess risk therefore, the area was not identified for sampling during Phase IIB.

The 2000 investigation was conducted to determine if the conditions of the Site 2 area were detrimental to the benthic community, since conditions may have changed due to three named storms and the natural movement of sediments. Station KL-34 was selected to be a delineating station to access the impact of the migration of these sediments. Survival rates at this station were 97% for the infaunal amphipod and 95% for the mysid shrimp, indicating no impact due to BEHP.