N00204.AR.002486
NAS PENSACOLA
5090.3a

COMMENTS FROM U S NAVY IN RESPONSE TO FLORIDA DEPARTMENT OF ENVIRONMENTAL PROTECTION COMMENTS TO FINAL REMEDIAL INVESTIGATION REPORT ADDENDUM 2 SITE 2 NAS PENSACOLA FL
5/9/2003
U S NAVY

Navy Response to Florida DEP Comments on the Final Remedial Investigation Report Addendum Site 2 Waterfront Sediments, NASPensacola
May 9, 2003

Comments received 15 May 03

Comment 1:
Tables 3-1 and 3-2: These tables have incorrectly reported reference concentrations by adding the two reference concentrations together. Reference concentrations for sediments should either be the average of the concentrations found in the reference samples or the range of the sample concentrations detected.

Response:
Section 3.1.1 of the text details how the data collected at the reference stations were used to develop site specific reference concentrations for individual constituents for comparison to Site 2 detections. Two times the mean concentration for each detection at Stations 18 and 22 was used as the reference concentration for each given constituent. This method and approach was agreed upon by the partnering team during the early stages of the DQO process.

Comment 2:
Section 3: Changing the reference concentrations will affect all discussions, tables and figures in Section 3.

Response:
The reference concentrations were calculated correctly, therefore no tables or figures need to be amended.

Comment 3:
Conclusions: A Feasibility Study needs to be discussed in this section.

Response:
A Feasibility Study will be developed for Site 2, as explained in Section 5.

Navy Response to University of Florida/Florida DEP Comments on the Final Remedial Investigation Report Addendum Site 2 Waterfront Sediments, NAS Pensacola
May 6, 2003

Comment 1:
At your request, we have reviewed the February 2003 *Final Remedial Investigation Report Addendum, Site 2 Waterfront Sediments, Naval Air Station Pensacola, Florida*. We reviewed a previous version of this document in a letter sent to you on 4 February 2002. Most of our comments were addressed in the present version, and we believe this revised document provides useful information for the evaluation of this site.

Response:
The Navy agrees with this statement.

Comment 2:
Our previous comments objected to the comparison of 1996 and 2000 sediment concentration data used to assert that concentrations are decreasing over time. We thought this was inappropriate because 1996 data were obtained from discrete samples, whereas information from 2000 was based on composite samples.

Response:
The Navy removed all comparisons between the 1996 and 2000 data. However, a summary of the 1996 data is provided in Section 1 as part of the rationale for the 2000 study.

Comment 3:
We also warned against relying on available sulfide information to predict future availability of metals. The current document does not include the concentration comparison and includes sulfides data only as ancillary information.

Response:
The Navy recognizes the AVS/SEM Model as a tool which can be used to enhance a data set and provide a theoretical measure of potential bioavailability for five divalent metals. When used in concert with other theoretical tools the AVS/SEM Model can be utilized in the weight of evidence approach.

Section 3.1.5 provides the AVS/SEM results in surface sediments. The data was calculated using two USEPA approaches. The Navy showed both outcomes in Table 3-4, and placed more emphasis in the USEPA 2001 method (versus the USEPA 1991 method) which correlated more with the other technical findings.

Comment 4:
As before, we think that the lack of overt toxicity observed on the bioassays coupled with the absence of significant effects on species diversity and abundance observed in the benthic community surveys demonstrate that contaminants present at the site are not having a significant adverse effect on benthic organisms, even though some Effects Range Medium (ERM) and Probable Effect Levels (PELs) are exceeded in some of the quadrantsstudied.

Response:
The Navy agrees.

Comment 5:
Given that none of the contaminants present at the site are expected to bioaccumulate significantly, the lack of direct effects suggests there are no significant adverse environmental effects due to Site 2 contaminants.

Response:
The Navy agrees.

Navy Response to US EPA Comments on the Final Remedial Investigation Report Addendum Site 2 Waterfront Sediments, NASPensacola April 30, 2003

Comment 1:
The conclusion section in this report states, "the multiple lines of evidence gathered during the investigation of Site 2 concluded that the area is recovering from past Naval Base Activities". Although the data may demonstrate that there is change in site conditions, it does not necessarily support a recovering effect.

Response:
The Navy had based their conclusions on the data presented (chemical analyses, toxicological exposures using two sensitive invertebrate species, and a benthic community survey). Table 4-8 summarizes the study results as agreed by the Partnering Team and captured in the DQO Process. Table 4-12 presents a weight of evidence table which also provides ancillary information to enhance the data and the Navy's conclusions. Toxicological exposures and benthic community surveys conducted in 1996 demonstrated signs of a degraded environment, whereas similar studies utilizing more sensitive species in 2000 demonstrated a healthy diverse community, which was interpreted as an improvement.

Comment 2:
The contaminants appear to have shifted over time from natural phenomena or normal dispersion; this would support a change in site conditions more so than a recovering effect.

Response:
The Navy, as a member of the Partnering Team, chose the Site 2 study location based in part on Phase IIB chronic toxicity test data collected in 1996. Sediments, especially those found at Site 2, are known to be in a dynamic (ever changing) system and can be expected to behave in response to weather and wave related conditions. The Partnering Team discussed at length the obvious weather conditions (three named hurricanes, and several winter storms) and the concept that sediments may have shifted. The concept of attenuation was also briefly discussed. The Team unanimously agreed to conduct their investigation based on the data collected during Phase IIB, and to concentrate their efforts in areas exhibiting HI values greater that 10. The problem statement agreed upon by the Team was, "Are chemicals in the Site 2 sediments creating a condition adverse to benthic communities?" Utilizing the DQO Process, the 2000 Site 2 Investigation developed an approach to answer this question.

Comment 3:
Additionally, the comparison of data from the two different sampling events can be performed on a generalized basis; it cannot be performed as an exact comparison.

Response:
The Navy agreed that direct comparison would not be applicable and removed direct comparisons between the 1996 and 2000 data. However, a summary of the 1996 data is provided in Section 1 as part of the rationale for the 2000 study.

Comment 4:
The last sampling event used the DQO process as a design standard which produced a more comprehensive sampling scheme then the initial sampling event. However, the data does not support the present day conditions of the site and a decision can be made in the next step.

Response:

An FS will be completed to evaluate appropriate alternatives.

Comment 5:
Figure 4-1, "Decision Flow for Each Decision Unit" states, if condition "1" or "6" of the triad exist in the top 6" of sediment, declare unacceptable condition, calculate remedial goal objectives and go to FS. The document identifies two decision units (DU) that demonstrate condition "6", CD-23 and EF-45. Also, NOAA's comments, EF-23 may be an additional area of concern. The next step in the process would be to calculate a remedial goal and proceed to an FS to evaluate alternatives.

Response:
Agreed, the Navy will develop a Feasibility Study to evaluate alternatives. Decision Unit EF-23 will not be included in the FS because when data was applied to the SQT, Condition 3 was interpreted as contaminants are not bioavailable.

Comment 6:
Keep in mind, a physical action will not necessarily be required, however, all alternatives should be evaluated and the most appropriate alternative selected.

Response:
The Navy agrees. The FS will evaluate all feasible remedial alternatives for the site, including a no-action alternative.

Editorial Comment # 1: Page 1-4, 2nd paragraph, 6th sentence is incomplete

Response:
Noted. The sentence should read: "HQs are produced by dividing the detected concentration by the sediment screening value (SSV) for that constituent."

Editorial Comment # 2: Page 4-19, 5th paragraph, last sentence - Tables 3-10 and 3-11 should be 3-8 and 3-9.

Response:
Agreed, the sentence will be revised.

Navy Response to NOAA Comments on the Final Remedial Investigation Report Addendum Site 2 Waterfront Sediments, NASPensacola
April 7, 2003

Comment 1:
Over a year ago, a similar conclusion appeared in a previous version of the subject report and NOAA, at that time, had the same comments (see NOAA memo dated January 23, 2001). The Navy's response to the NOAA comment was that previous (1996) sampling would be deleted and the revised document would just focus on the most current (2000) dataset.

Response:
The Navy reviewed the comments received from all the Stakeholders and Trustees and agreed to remove direct comparisons between the 1996 and 2000 data. However, a summary of the 1996 data is provided in Section 1 with the intention of providing past data to support the level of investigation undertaken in the 2000 study so informed decisions about the site can be made.

Comment 2:
The Navy response implied that the recovery issue would not be addressed. Indeed, the subject document never technically evaluates recovery at Site 2. Therefore, the "recovery" conclusion is not warranted and should be deleted.

Response:
The Navy's response to the 'Draft Remedial Investigation Report Addendum Site 2 Waterfront Sediment 2002' comments indicated that "...narrative portions of the report will be revised to provide supplemental technical support to the 2000 findings". The 2000 Study was focused on the following question, 'Are chemicals in Site 2 sediments creating a condition adverse to benthic communities?' Based on the study conducted in 2000, conditions have improved when comparing toxicological and species diversity data collected in 1996. The Navy is aware that direct comparisons between 1996 and 2000 data cannot be conducted due to different sampling strategies, sampling design, and test endpoints. However, the Navy focused attention on the site based on the 1996 results, and improved their approach. The 2000 Site 2 sampling effort found that although some chemicals still exist in Site 2 sediments, they are not affecting the benthic community. In fact, the benthic community was found to be diverse and abundant at the site.

Chemicals do not seem to be adversely effecting sediments (with the exception of the 2 units moving on to a FS) when compared to site specific reference locations.

Comment 3:
Rather than "recovering", multiple lines of evidence (see comments below) suggest the system has probably reached some level of equilibrium.

Response:
The Navy conducted a thorough study which incorporated a sound sampling design. The data was then analyzed so responses to questions asked in the DQO document could be answered. Data collected in 2000 showed conditions to be improving, when compared to the earlier 1996 study. High diversity, richness, and species abundance data, coupled with the acute and chronic responses of sensitive invertebrates exposed to Site 2 sediments were the basis for the Navy's conclusion that the sediments in that area are 'improving'. Measuring improvement or recovery was not one of our

original goals in this sampling and as a result the improved conditions could be due to other factors i.e., seasonal, methodology, etc.).

Comment 4:
Untreated industrial wastes were routinely discharged into Pensacola Bay at Site 2 for more than 30 years from 1939 to 1973 (§1.1).

Response:
The Navy included this information as background and to help focus the investigation as to what type of constituents might be found in the area sediments. Based on this information, chemical constituent analyses focused on inorganics, organics, pesticides, PCBs, low level PAHs, and organotins (Section 3.2.2 DQO Plan).

Comment 5:
After industrial wastes were redirected to the base WWTP in 1973 (§1.1), Site 2 has had sufficient time (30 years) to reach some level of equilibrium.

Response:
The Navy understands the concept of 'equilibrium' and speculates that if the source is removed (untreated industrial wastes from 1939 to 1973) and clean water flows over the area once a day (through tidal action) for 28 years, with the occasional hurricane disturbing the sediments, equilibrium could occur.

Comment 6:
Hydrodynamics of Site 2 (i.e., siltation and rotational flow patterns described in §2.1) suggest a more or less self-contained hydrologic unit. These site characteristics would favor retention of contaminants released into the immediate environment. Indeed, the observation of paint chips in Site 2 sediments by divers and laboratory personnel suggest industrial wastes discharged into Pensacola Bay remain.

Response:
Prior to the commencement of the field component of the Site 2 study, investigators speculated an 'eddy effect' was occurring based on the silty sediments collected in 1992 and 1996 near the seawall. However, it wasn't until the USEPA divers were actually collecting the sediment samples that an underwater seawall was discovered. The seawall undoubtedly explains the 'eddy effect' and the concept of a self-contained hydrologic unit. Under day-to-day weather patterns sediments may be confined to the area, but it could also be speculated that conditions change during winter storms, high rainfall (surface water runoff from the well documented stormwater outfalls which line the shoreline), and tropical disturbances which may create a dispersion of these silty sediments outside the submerged seawall.

Comment 7:
Site-wide sediment chemistry in 1996 is similar to that observed in 2000. (See NOAA January 23, 2001 comments for detail). At that time, NOAA suggested a simple procedure for quantifying temporal differences. This procedure has apparently not been pursued.

Response:
The Navy received comments from both the Stakeholders and the Trustees indicating that there were too many uncertainties connected to a comparison of the 1996 and 2000 data (see 2003 UF Comment 2, and USEPA Comment 3). NOAA, as a Trustee, provided a method by which the data could be 'normalized' (to quantify temporal differences) to salvage the data and make comparisons possible. The Navy chose to summarize the 1996 data in Section 1, and base their study on the problem statement developed by the Team (in which NOAA was an active participant) which was, "Are chemicals in the Site 2 sediments creating a condition adverse to benthic communities?" Utilizing the DQO Process, the 2000 Site 2 Investigation developed an approach to answer this question and base their response on the data collected during the 2000 study. A summary of the 1996 data is provided in Section 1 to support the level of investigation undertaken in the 2000 study and to provide previously collected information so informed decisions can be made.

Comment 8:
Significant sediment toxicity was observed in two separate bioassays in both 1996 (Table 1-2) as well as 2000 (Tables 3-8 and 3-9). The main area of toxicity and elevated chemistry appear to have shifted slightly eastward between 1996 and 2000. This eastward shift was an important factor in selecting sample locations for the 2000 field study.

Response:
The Navy agrees that there was significant (statistical and biological) toxicity observed in the 1996 study, and assents that statistical significance was identified in 18% of the *Leptocheirus* bioassays, and 0% of the mysid exposures (when viewing quantal endpoints) conducted in the 2000 Site 2 study. The Navy asserts that these low effect values do not appear to be biologically significant, especially when no concurrence can be correlated between quantal and non-quantal effects at each station.

The Navy focused their 2000 investigation on the area which exhibited the most impact in 1996 (Table 1-2), and provided a synopsis of the 1996 data set to glean information for the 2000 study and make informed decisions on the new data set possible.

Comment 9:
Indices measuring the benthic macroinvertebrates community are similar in 1996 (Table 1-3) and 2000 (3-10). As noted in NOAA's 2001 comments, "the small differences observed may simply reflect normal seasonal variability. The 1996 samples were collected in winter (January) while the 2001 samples were collected during the spring biological recruitment period (March). 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:
The report briefly discussed seasonal variability and providedtemperature conditions for the area waters during the 1996 and 2000 sampling effort. An average of the 1996 benthic diversity indices (using only those stations common to the 1996 and 2000 sampling points) exhibited an average diversity of 2.26, average evenness of 0.69, and average richness of 5.68. An average of the 2000 benthic diversity indices exhibited diversity to be 3.08, evenness to be 0.81, and richness to be 8.66. The Navy does not view the increase in diversity as 'similar'; the Navy translates this phenomena as improved site conditions. In addition, benthic diversity was considered to be a 'snapshot' of the current conditions of the site (as evidence for the sediment quality triad) and not a benthic community study in which several years would be committed to studying the species recruited seasonally to the area. Again the Navy's conclusion on this site is based on the 2000 data set only and not a comparison between two separate events.

Comment 10:
The Navy has not followed it's own DQO Process for evaluating and managing "unacceptable conditions" at Site 2 Decision Units (DU).

The Navy went to great lengths to develop a Data Quality Objectives (DQO) Process for the2000 field investigation. Per Figure 4-1, that process specifies that if, "Condition 1 or 6 of the triad exists in the top 6" of sediment" then one must : 1) "Declare unacceptable conditions exist in the DU", 2) "Calculate Remedial Goal Objectives" and 3) "Go to a Feasibility Study". The subject report concluded that DUs
EF-45 and CD-23 were both Condition 6 (§5.0). The report, therefore, should follow the Navy's DQO Process and conclude that conditions in these two areas of Site 2 are "unacceptable", calculate RGOs and proceed to an FS. It is interesting to note that EF-45 and CD-23 are located next to one another in the eastern portion of Site 2 (Figure 4-17). This eastern area should be the focus of the FS for Site 2.

Response:
The Navy has recommended a Feasibility Study for Site 2, which will focus on Decision Units EF-45 and CD-23, as stated in Section 5.

Comment 11:
Delete §2.2.3 "USEPA Sediment Toxicity Criteria"
After conversations with EPA Region 4 senior ecological risk assessor (Lynn Wellman), NOAA has concluded that this section erroneously portrays an EPA document as recommending "Sediment Toxicity Criteria". That document, EPA 1991b, (Which is incompletely cited in Reference §6.0) provides national technical support for evaluating aquatic effluents, not sediments. Moreover, the toxicity threshold suggested in this section, 25%, contradicts the 20% value previously accepted by the NAS PNS team. Finally, the section is internally inconsistent. The second sentence suggests the 25% effluent criterion should be applied only to non-quantal sublethal responses such as growth or reproduction while the forth sentence indicates it's appropriate only for the lethality endpoint.

Response:
The Navy appreciates the clarification and understands NOAA's position. The Navy's evaluation of the toxicity data used the agreed upon 80 % survival.

Comment 12:
Conduct statistical analysis of mysid reproduction data as previously promised by the Navy.

A. Over a year ago, NOAA suggested the mysid reproduction data be statistically analyzed as was done for all bioassay endpoints (see January 23, 2001 NOAA memo). The Navy concurred with this comment and agreed to statistically analyze the mysid reproduction data and report results in the main body of the revised document. This has not been done (Table 39 and §3.2.3).

Response:
The Navy responded to the earlier comment by agreeing to include the statistically analyzed data, hence the coefficient of variation (%CV) column added to Tables 3-8 and 3-9. The CV incorporates the standard deviation relative to the mean, and at the same time provides information on the intra-test variability. Both Tables 3-8 and 3-9 show low variability, indicating the tests were run by an experienced facility.

Appendix C has a compilation of the statistical analyses which were performed to satisfy comments from NOAA in 2002. CV's were included in the tables and a brief explanation provided in Sections 3.2.2 and 3.2.3.

Comment 13:
B. Mysid reproduction appears adversely affected in two of the three C-17 Control treatments (Table 3-9). Therefore, statistical comparisons to this treatment appear problematic. In the first and third control tests only 50% of females produced eggs while over 90% of females in the second control produced eggs. Low mysid reproduction may be related to the very high silt content (99%) of C-17 sediments relative to all other sediments tested. For example, reference Station 22 had the second highest silt content, 61% and very low rates of reproduction; only 22% of females produced eggs. Station 18, the reference location closest to Site 2, had silt content (42%) approximating Site 2 sediments and had much higher rates of mysid reproduction; 93% of the females produced eggs. Consequently, reference Station 18 should be the sole statistical basis for analyzingthe effects of Site 2 sediments on mysid reproduction.

Response:
The Navy would like to emphasize that the test acceptability criteria (TAC) for the mysid chronic test is that 50% of the females in the control exposure must be gravid if the test series is to count as a valid test. The Site 2 sediment controls used by the laboratory (C-17) met the TAC established for these long term exposures. USEPA has dropped fecundity as an endpoint from the promulgated methods due to the high variability experienced in both inter- and intra-laboratory exposures. However, the Navy agreed to include the fecundity endpoint to enhance the data and facilitate informed decisions.

The laboratory which conducted the tests informed the Navy that the C-17 sediments used for these exposures and the *Leptocheirus* exposures came from the same batch and should be uniform in

their physical composition. An explanation of the high variability between control fecundity may simply be the individual genetic differences between organisms, which may be why USEPA made this endpoint optional as of the 2002 rule.

As explained in Section 4-7, TAC for 7-day mysid chronic tests include: 80% or greater survival, average dry weight of 0.20 mg or greater in the controls, and fecundity may be used if 50% or more of females in controls produce eggs. TAC for 10-day Leptocheirus acute test includes: minimum mean control survival of 90% in the control exposure. Organism weight at test termination will be compared to the control exposures and calculated using a Ttest.

The table below combines all the bioassay endpoints and identifies statistically significant differences between the Reference Stations (a and b) and the Negative Controls - C-17 (c, d, and e). This information is also presented in Section 4-7 text.

Station	Mysid Survival	Lepto Survival	Mysid Weight	Lepto Weight	Mysid % ♀/eggs
IJ-12	100 [a]	89 [a]	0.36	0.13	85
GH-12	100 [a]	88 [a,b]	0.34	0.14	92
IJ-34	97	95	0.39 [a]	0.25 [a,b]	94
GH-34	97	95	0.35	0.21 [a,b]	100
EF-23	97	95	0.34	0.16 [a]	67
IJ-56	97	81 [a,b]	0.53 [a,b]	0.21	94
GH-56	97	82 [a,b]	0.4	0.28 [a,b]	88
EF-45	97	73 [a,b]	0.37	0.24 [a,b]	75
CD-23	100 [a]	78 [a,b]	0.47 [a,b]	0.27 [a,b]	95
KL-34	95	97	0.34	0.27 [a,b]	88
GH-67	84	99	0.30	0.13	69
Ref Station 18 [a]	90	98	0.28	0.12	93
Ref Station 22 [b]	92	96	0.32	0.14	22
C-17 (# 1) [c]	97	93	0.32	0.22	53
C-17 (# 2) [d]	100	94	0.40	0.22	93
C-17 (# 3) [e]	93	100	0.32	0.23	53

Lastly, the fecundity endpoint was added to enhance the data collected during this investigation, since a decline in reproduction could affect the food web in this general area. USEPA made this endpoint optional as of the 2002 rule when the chronic manual was updated in 2001, due to the variability demonstrated between laboratories.

Comment 14:
C. When appropriate statistical analyses have been completed for the mysid bioassay, it may be found that "unacceptable conditions" exist at additional areas within Site 2. For example, only 67% of female mysids exposed to sediments from EF-23 had eggs. EF-23 is adjacent to the two co-located DUs (EF-45 and CD-23) where "unacceptable conditions" were found (§5.0). Therefore, if conditions at EF-23 are found to be "unacceptable", the eastern area of Site 2 would enlarge for the FS (Figure 4-17).

Response:
Mysid conditions at Stations EF-45 and CD-23 were acceptable with 97% and 100% survival, 75% and 95% females with eggs, and measurable weight gain. Leptocheirus exposures revealed a toxicity response (statistically significant) at the same stations, but no effect in the sub-lethal (non-quantal) growth endpoint. Station EF-23 had 97% mysid survival, 67% females were gravid, and all animals in this sediment experienced, measurable weight gain; the Navy did not see a negative response from mysids at this station. Station EF-23 also had 95% survival results with Leptocheirus, and adequate growth. Based on these findings and translated into the SQT, Condition 3 exists. The Navy sees no basis in adding EF-23 to those considered unacceptable.

Comment 15:
The bioavailability of inorganics is a significant uncertainty for Site 2 COPCs that has been virtually ignored in the subject report.

A. Visual observations of paint chips in Site 2 sediments suggest industrial wastes discharged into Pensacola Bay for 30+ years still remain. If paint chips are visible to the naked eye (see Table 2-1), it is likely that smaller, less visible chips are present throughout Site 2. If one assumes inorganics in paint chips are less available, this may help explain some of the results and the "false-positives" cited in the conclusions (§5.0, ¶2).

Response:
Divers noted glass eels and paint chips at Stations IJ-12 and GH-56, and submerged aquatic vegetation in the shallow areas of IJ-12 (Table 2-1). Table 3-1 lists several metals, high molecular weight PAHs, and PCBs at IJ-12, which is adjacent to a seawall stormwater outfall. Station GH-56 was found to have one exceedance, naphthalene (62 ppb), which was commonly found throughout the site. Mysid survival at these stations was 100% at IJ-12 and 97% at GH-56. Growth at GH-56 was normal when compared to the laboratory control C-17 (Table 3-9), and out-performed the USEPA Reference Stations. Fecundity at both IJ-12 and GH-56 was very high, with 85% and 88% (respectively) of the females gravid. *Leptocheirus* survival at Station IJ-12 was 89% and 82% at GH-56, with growth found to be statistically significant at Station IJ-12, and normal at GH-56. Growth was also found to be statistically significant at the two USEPA Reference Stations. The presence of paint chips may lead to high chemistry that would indicate effects should be seen; however, the physical states of the contaminants (bound up in visible paint chips) do not produce the expected results (lack of toxicity). Based on these observations (presence of glass eels, and vegetation), high mysid survival, good *Leptocheirus* survival, few sublethal effects, and agreed upon endpoints, the presence of paint chips does not appear to be adversely affecting the benthic community.

Comment 16:
B. 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 sufficiently bind all 5 divalent metals 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 expressed.

Response:
The Navy agrees with the theory that SEM/AVS can be used to determine if trace metals may be bioavailable to benthic organisms (Section 3.1.5). The Navy recognizes that when SEM-AVS values exceed 1, as they do at Station CD-23, and are complimented by the SEM/AVS value >1, the theory is more plausible, and the benefit of this tool is appreciated.

Comment 17:
C. Consistent with theory, when SEM:AVS ratios exceed 1, as they do at CD-23, GH-67, and GH-12 (Table 3-4), amphipod growth or survival is significantly diminished (Table 3-8). CD-23 was one of the two DUs at Site 2 where "unacceptable conditions" were found.

Response:
The Navy understands the use of these tools and the uncertainty surrounding them. As such, the Navy did not put the weight on their usefulness; but rather used them as an ancillary tool to enhance the study's findings. In addition, UF risk assessors voiced their concerns about relying on the SEM:AVS data too heavily.

For instance:

Station	Mysid (Survival)	Mysid (Growth)	Mysid (Fecundity)	Leptocheirus (Survival)	Leptocheirus (Growth)	SEM/AVS (USEPA 1991 Method)	SEM-AVS (USEPA 2001 Method)	Diversity
CD-23	100%	0.47	95% with eggs	78%	0.27	2.4	1.67	3.04
GH-67	84%	0.3	69% with eggs	99%	0.13*	1.6	0.043	3.06
GH-12	100%	0.34	92% with eggs	88%	0.14*	1.30	0.095	2.90

The Table above was recreated from data presented in Tables 3-8, 3-9, 3-4, and 3-10. Concurrence between species and endpoints does not correlate to SEM/AVS or SEM-AVS findings, and has not affected the diversity. Further, data from Table 3-1 indicates that only Station CD-23 had Cd, Cu, and Pb levels above sediment benchmark values. Stations GH-12 and GH-67 did not have any inorganic exceedances. Based on these inconclusive findings, the Navy does not put much significance on this tool.

Comment 18:
§2.1 In this section, report the characteristics of the two reference locations (18, 22) shown in Figure 1-1 (e.g., depth, distance from Site 2, sediment grain size) and the rationale for their selection.

Response:
Page 11 of Appendix A (Data Quality Objectives Summary, Work Plan, and Sampling and Analysis Plan - which NOAA contributed to as a DQO Planning Team Member) provides the following information: "Reference stations should emulate the decision units of the site with regard to grain size, chemistry and toxicity. Therefore, it is desirable to select two reference stations, one with approximately 20% sand content and one with approximately 80% sand content, as sand content is the common denominator. Lower Pensacola Bay areas might be suitable back-up reference stations if no others can be identified.

USEPA Pensacola Bay Stations 18 and 22 were selected as the reference stations for Site 2 based on similar sand (%) components, high amphipod survival rates when exposed to sediments for 10 days, and healthy benthic indices in past USEPA studies (1992 & 1996). The average depth of Station 18 is twice that expected at Site 2, but phone conversations with several benthic ecologists (Gary Gaston, University of Mississippi, Richard Heard, Gulf Coast Research Laboratory, Tony Martin, Barry Vittor and Associates, and Virginia Engle, USEPA EMAP Coordinator) indicated that the fauna in this shallow bay system would be similar, and that sand, silt, and clay are the factors that drive habitat recruitment and not depth." This text will be added to the document.

Comment 19:
§3.1.6The method to estimate Total Organic carbon (TOC) requires additional clarification and substantiation. The test says a sample from each DU was analyzed for "organic and non-organic % for material than 2mm". This reviewer does not understand what the text is saying. TOC was later estimated (from what it's not clear) using an approach based on a personal communication with a California Contractor. This approach requires better substantiation. Is there a published laboratory method or even an internal lab SOP that can be cited?

Response:
The Navy entrusted the sampling and analysis portion of this study to be conducted by USEPA. Several unforeseeable situations arose in which subsurface sediment samples could not be collected due to refusal (thus providing some data gaps) and in the case of TOC (a vital component of sediment analyses), the samples were lost.

In order to salvage any portion of the data, the Navy took the conservative approach of estimating that 50% of the organic material caught on a sieve (<2mm in size) may constitute a similar value of TOC found in area sediments. This conceptual approach was based on a lively discussion held at a sediment workshop, attended by sediment experts.

Comment 20:
§3.2.1The negative "control" sediment, C-17, is not a true control for sediment testing (see NOAA comments dated January 23, 2001). Rather it's an additional reference sediment located in Perdido Bay just west of Pensacola Bay (Figure 1-1). Because C-17 sediments have a very high silt content (99%) they are physical dissimilar to Site 2 sediments (see Tables 3-5 and 3-6). C-17 sediment chemistry results in Appendix C are incomplete and have limited value to the Site 2 bioassays. That is, 1) no inorganics data are reported (a major class of COPCs at Site 2) and 2) it is likely that these chemistry data were not collected synoptically with Site 2 sediment bioassay samples.

Response:
C-17 is a true negative sediment control which has been used since 1990 by the USEPA and NOAA for their sediment studies including the EMAP Program and the National Status and Trends Program. Aliquot samples are analyzed yearly, and a synopsis is included in Appendix C. Organism responses in C-17 have been consistently good as supported by the continued use by these national programs.

The purpose of a negative sediment control is to have a basis by which to compare the results of 'unknown' sediments being tested (performance standard). Negative controls are routinely run with each set of tests, and are handled in the same manner as the sediments being tested. The laboratory typically uses a sediment which has been tested routinely, with different organisms,in which consistent results (low CV's) can be obtained. USEPA Environmental Research Laboratory has furnished the chemical analyses for several years since the contract laboratory routinely does the USEPA's sediment testing.

For clarification purposes, a positive control is also known as a reference toxicant test in which laboratory organisms are challenged against doses of a known quantity to ascertain the organisms sensitivity by causing effects (i.e., mortality). Positive control tests are typically run monthly and data is presented in the form of a control chart.

Comment 21:
Define "TAC" in §3.2.1. Also, an incorrect figure is cited for locating C-17 in the report.

Response:
Pages iv-v lists abbreviations and defines TAC as test acceptability criteria.
Figure 1-1 correctly identifies the location where C-17 is collected. Section 3.2.1 will have the Figure number corrected.

Comment 22:
§3.2.2 Delete last sentence in the first paragraph as well as the entire third paragraph per comment 3 (Comment 11) above. Consequently, one can delete % CV columns in Tables 38 and 3-9.

Response:
C-17 data was included in the Final Remedial Investigation Report Addendum based on comments received from NOAA January 2002. In addition, NOAA alluded to concerns with control performance, asked that normalized data be presented in a 'non-normalized format', and that statistical significance of sublethal endpoints be expressly compared to both reference stations. The

Navy complied by showing the negative control (C-17) performance, adding a column for the % CV, and comparing the various endpoints of each test against the Reference Stations 18 and 22, and exhibiting this information on Tables 3-8 and 3-9 with superscripts.

The CV incorporates the standard deviation relative to the mean, and at the same time provides information on the intra-test variability. Both Tables 3-8 and 3-9 show low variability, indicating the tests were run by an experienced facility.

Comment 23:
Tables 3-1 and 3-2 Surface and subsurface chemistry data are erroneously compared to Site 40 Bayou Grande data (see footnote 'd', EnSafe 1999). Therefore, eliminate this comparison and delete the "Reference Conc." column from both tables. Comparison should be to the appropriate reference location; probably Station 18 due to its proximity. For the subsurface sediment table, report the depth interval represented by individual samples.

Response:
Section 3.1.1 details how the data collected at the reference stations were used to develop site specific reference concentrations for individual constituents on the base for comparison to Site 2 detections. Two times the mean concentration for each detection at Stations 18 and 22 was used as the reference concentration for a given constituent.

The footnote reference refers to the basewide levels for DDT, DDE, and DDD, which the team agreed to use when reviewing data for these constituents. Since Site 2 is part of the base, it is appropriate to maintain the accepted basewide levels as a comparison.

The Navy disagrees concerning the removal of the "Reference Concentration" column from the tables, since DU constituent comparisons were made to sediment benchmark values and the Site 2 Reference Concentration.

Comment 24:
Tables 3-8 and 3-9 In previous comments (January 23, 2001) NOAA requested a variance term (SD or SE) be reported for each value in the amphipod and mysid sediment bioassays, including controls. The Navy's Response to Comments indicated this would be done for the revised report. It was not.

Response:
The Navy included the CV (a standard statistical measure of the relative variation of a distribution or set of data, defined as the standard deviation divided by the mean) to meet the intent of this request.

Comment 25:
In Table 3-8, it is unclear how treatments denoted with 3 asterisks differ from the other treatments. Why not simply report number of replicates (n=x) for each treatment in both tables.

Response:
The Navy felt using foot notes in the table was the clearest way to present this information. The foot note that applies this (***) represents four replicates of 20 organisms were exposed to sediments giving a total of 80 animals/sediment.

Comment 26:
Should "Criteria" in Table 3-9 be "Notes" or "Footnotes"? Superscripts a and b are not explained in this table.

Response:
The Navy combined the 'Notes' portion of the table with the added reference information which supplies the necessary criteria for the chronic endpoints of this test. The Navy will add the definition of the two superscripts used on Table 3-9.

Comment 27:
Table 3-10 In previous comments (January 23, 2001), NOAA requested clarification regarding what is meant by "* = Stations with species diversity overlap". Does it mean that stations with asterisks are different from other stations, or not different from each other, or what? The Navy's RTC indicated a "small diagram" would be provided explaining the "shared diversity data". I cannot find this diagram or an explanation that would enlighten the reader.

Response:
Section 3.2.4 states the following:
"Table 3-10 presents the benthic community species diversity results. The sampling plan was developed so that each DU was adjacent to the next (Figure 2-1), allowing for benthic community samples to be shared between Stations GH-12 and IJ-34; GH-34 and IJ-56; and CD-23 and EF-45 (these DUs are identified with an asterisk in Table 3-10)."

In the 'Notes' portion of Table 3-10, an asterisk is followed with the following explanation: "Stations with species diversity overlap".

Comment 28:
Figure 3-21 Delete "NO GROWTH OR FECUNDITY EFFECTS OBSERVED" in this figure. Sediments at EF-45 had significant adverse effects on mysid growth (Table 3-9) and, per comment 4 (Comment 12) above, no statistical analysis of mysid reproduction has been reported.

Response:
The Navy understands that the statistical program used to analyze the results of these tests 'flagged' this sediment as indicating a statistically significant sub-lethal effect. Mysid results for sediments collected at Station EF-45 indicated 97% survival, 75% of the females were found to be gravid, and weight gain easily surpassed the two USEPA Reference Stations. The definition (which was inadvertently omitted, NOAA Comment 26) for the superscript provided adjacent to the weight value in EF-45 shows a statistical difference was noted when compared to Reference Station 18. Further review of this significance clearly shows that the sediment at EF-45 out performed the Station 18 sediment. The notation, 'In relation to the Reference Station Sediment' will be added.

Comment 29:
§4.3 Delete first sentence suggesting benthic community diversity has increased between 1996 and 2000. As noted above an in previous NOAA comments, differences in season and collection methods may well explain any differences observed.

Response:
The report briefly discussed seasonal variability and providedtemperature conditions for the area waters during the 1996 and 2000 sampling effort. An average of the 2000 benthic diversity indices exhibited diversity to be 3.08, evenness to be 0.81, and richness to be 8.66. Based on the 2000 results, benthic conditions have improved as evidenced by the diversity index and outcomes.

Comment 30:
Table 4-6 Per results in Table 3-9, add a "+" for adverse effects on mysid growth at Station EF-45.

Response:
The Navy understands that the statistical program used to analyze the results of these tests 'flagged' this sediment as indicating a statistically significant sub-lethal effect. However, actual mysid results for sediments collected at Station EF-45 indicated 97% survival, 75% of the females were found to be gravid, and weight gain easily surpassed the two USEPA Reference Stations. The definition (which was inadvertently omitted, NOAA Comment 26) for the superscript provided adjacent to the weight value in EF-45 shows a statistical difference was noted when compared to Reference Station 18. Further review of this significance clearly shows that the sediment at EF-45 out performed the Station 18 sediment. The definition will be added to Table 3-9 for clarification.

Comment 31:
Page 4-21, in the third full paragraph, delete phrase "however, this effect did not carry over to the sublethal endpoints". Just because significant amphipod mortality was observed at EF-45 does not necessarily mean sublethal effects will be forthcoming (and vice versa). Delete or revise last sentence in this paragraph which states "mysid shrimp test performed very well (all endpoints) when exposed to the same sediments". Mysid growth was significantly depressed in one treatment and mysid reproduction has yet to be analyzed. References to tables are incorrect in this paragraph.

Response:

Agreed, the Navy will delete the phrase.

Comment 32:
In the first full paragraph, delete the phrase "however, this effect did not carry over to the sublethal endpoints". Just because significant amphipod mortality was observed at CD-23 does not necessarily mean sublethal effects will be forthcoming (and *vice versa*). Revise last sentence in this paragraph to more accurately describe Condition 6; i.e., unacceptable conditions exist at CD-23. References to tables are incorrect in this paragraph.

Response:
Agreed, the Navy will delete the phrase.

Comment 33:
Table 4-9 For the subsurface sediment table, report the depth interval represented by individual samples.

Response:
Section 2.1.1 describes sample collection and discusses the purpose of the sediment core collection as follows: "The intent of the sampling was to collect surface (0"-6") and subsurface (6"-36") sediments at each location to identify the vertical and horizontal extent of contamination. Sample collection techniques are described in the Final report, Pensacola Naval Air Station, Sediment Survey, Operable Unit 3 (USEPA, 2001) and are briefly described below." Because of compaction of the cores, the actual depth interval is difficult to determine.

Comment 34:
Table 4-11 In this table comparing surface and subsurface sediment chemistry, delete rows where comparisons are not possible; i.e., no subsurface samples at GH-12, IF-56, GH-56, EF-45, GH-67. A critical data gap is the lack of subsurface chemistry for EF-45 where "unacceptable conditions" exist in surface sediments.

Response:
As explained earlier, the sediment core experienced refusal at several stations. The Navy does not see a benefit in removing the information from the table as they are clearly labeled as "Not Sampled".

Comment 35:
§5.0, ¶1 Text suggests the Navy is still releasing anthropogenic contaminants into Pensacola Bay in the vicinity of Site 2. If true, regulatory action may be warranted.

Response:
A Feasibility Study will be developed for Site 2. Stormwater discharge is handled under the Stormwater Management Program. Petroleum discharges are handled under the State's Petroleum Program.

Comment 36:
§5.0, ¶2 Only false-positives are discussed. A more balanced presentation would note that when SEM:AVS ratios exceed 1, as they do at CD-23, GH-67, and GH-12 (Table 3-4), amphipod growth or survival is significantly diminished (Table 3-8).

Response:
Please see the response to Comment Number 17.

Comment 37:
§5.0, ¶3 Again, more balanced interpretation of results is required. Conclusions state no DU was classified as Category 3 using mean ERM Quotient yet EF-23 was Category 3 using Florida DEP mean PEL Quotient (Table 4-5).

Response:
The Navy was not discussing PELs in §5.0, ¶3, As stated, the Navy was discussing ERM Quotients.

Comment 38:
§5.0, ¶4 Delete subjective characterizations such as "do not appear". For a more balance discussion, include the possible influence of depth, grain size and freshwater at the reference stations on benthic community structure.

Response:
The sentence will be modified to read, "Biologically, benthic organisms are being recruited to the Site 2 area. The species diversity was greater at Site 2 than...."

Depth, grain size, and freshwater (Does NOAA mean the lower salinity recorded at Reference Station 22?) at the reference stations on the benthic community were discussed in Sections 2.1.1 (page 2-4), 3.1.1 (page 3-1), 3.1.7 (page 3-30), and 3.2.4 (page 3-38).

Comment 39:
§5.0, ¶5 Delete the subjective and misleading characterizations which follow "...CD-23 were determined to be condition 6". Terms such as "marginal" and "high" are subjective. Concluding mysid reproduction was unaffected is misleading because that endpoint has yet to be statistically analyzed. Delete the last sentence. All statistical tests in the amphipod bioassay were compared to the "control" C-17 sediment, not to reference sediments.

Response:
The Navy agrees to modify responses and will exchange 'marginal' and 'high' with 'relate to' and 'with respect to' in order to remove the subjective terms.