UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
REGION IX
75 Hawthorne Street
San Francisco, Ca. 94105-3901

N00217.002738
HUNTERS POINT
SSIC NO. 5090.3

Raymond E. Ramos
Western Division
Naval Facility Engineering Command
900 Commodore Drive
San Bruno, CA 94066-2402

Subject: Environmental Sampling and Analysis Plan Data Review

Dear Mr. Ramos:

Enclosed please find the Environmental Protection Agency's (EPA) comments on the results of the Environmental Sampling and Analysis Plan (ESAP) which was implemented at the Hunters Point Annex (HPA). The ESAP data was submitted to the EPA on April 21, 1993. The submittal was data only--no interpretation of the results was included.

It is our understanding that as part of the Phase 1A Ecological Risk Assessment (ERA), to be presented to the agencies on March 18, 1994, the ESAP data will be more formally interpreted. Based on the analysis of the ESAP data and other existing environmental data, coupled with descriptions of on-site habitats and species of concern, the Navy will also on March 18, 1994, propose the second phase of the ERA--Phase 1B. Phase 1B will include additional field work, if necessary, to fill in data gaps identified as part of Phase 1A.

Regarding the ERA, EPA's comments on the ESAP data results are meant to help the Navy identify the data gaps which should be filled as part of Phase 1B. This is done by providing a preliminary interpretation of the data such that successful and failed data collection efforts are discerned.

Beyond the ERA, however, EPA's comments on the ESAP data results are meant to help the Navy identify physical areas which the results of the ESAP suggest require further investigation and perhaps remediation. That is, regarded as Site Investigation (SI)-level data, some of the ESAP data results suggest the need for further Remedial Investigation (RI)-level study. Whether these areas are formed as another operable unit, an extension of existing operable units, or some other organizing mechanism is a subject worthy of inter-agency discussion. We recommend that a meeting be held in the near future to discuss the results of the ESAP, agency comments, and next steps for action.

Printed on Recycled Paper

As a summary, the ESAP provides useful chemical data for the purpose of identifying areas of contamination in need of both further study and perhaps remediation. In particular, the near shore sediments appear to be significantly contaminated such that consideration should be given to potential remedial alternatives as a way of focussing future studies of this area. Such focus will hopefully prevent the unnecessary expenditure of time and resources on extraneous study such that the best method of remediating the near shore contamination can quickly be achieved.

In addition, the bioaccumulation data suggest that, though not evident in the limited water quality data, water column contamination has occurred such that aquatic organisms may be bioaccumulating those contaminants. This matter, too, should be more thoroughly studied and appropriate remedies pursued.

Further, while the bioassays were generally unsuccessful, they nonetheless offer some valuable insights regarding water and sediment toxicity. In particular, the toxicity data will be useful for focussing future surface water and sediment investigation and remediation. That is, the data should not simply be discarded. Nor, should the bioassays be re-performed as written. Instead, all of the ESAP sampling stations from which the bioassay and chemical data suggests the presence of toxicity or toxins should be identified for further study either under the Ecological Risk Assessment (ERA) or under a Remedial Investigation.

As a final matter, we suggest that a data correlation effort be undertaken to better compare the results of on-shore investigatory work with the near shore data results. It is apparent that contaminants in the near shore, at least in part, originate from the shore operations. As such, a correlation of the various Preliminary Assessment/Site Investigation (PA/SI) and Remedial Investigation/Feasibility Study (RI/FS) results with data otherwise meant to support the ERA, should help identify logical source reduction projects, potential removal actions, and remedial actions--where needed.

It should be noted that EPA's preliminary effort at interpreting data and correlating data can be considered a rough, first step, only. We performed no statistical analyses or investigations of our own. Instead, we endeavored to compare the ESAP data with known screening criteria and with mean reference area values, where reference area data were available.

Please note that EPA's Office of Research and Development has prepared comments related to the Navy's statistical analysis conducted on the bioassay test results. Those comments are included as an attachment to this letter.

Please also note that the National Oceanic and Atmospheric Administration (NOAA) has submitted comments to EPA regarding its review of the ESAP data. NOAA's letter is also included as an attachment and the comments fully supported by EPA.

If you have any questions regarding this review, please contact Alydda Mangelsdorf at (415) 744-2389 or Roberta Blank at (415) 744-2385.

Sincerely,

Alydda Mangelsdorf
Remedial Project Manager

Roberta Blank
Remedial Project Manager

Attachments:
1. EPA's ESAP Review Comments, including ORD's review of the Navy's statistical analyses
2. NOAA's ESAP Review Comments

cc:
Dave Song, WESDIV
Barbara Smith, RWQCB
Cyrus Shabahari, DTSC
Denise Klimas, NOAA
Gary Welshans, PRC
Amy Brownell, City and County of SF
Jim Sullivan, NSTI

ATTACHMENT 1

EPA'S ESAP REVIEW COMMENTS

ENVIRONMENTAL PROTECTION AGENCY
REGION IX
75 HAWTHORNE STREET
SAN FRANCISCO, CA 94105

TABLE OF CONTENTS

Executive Summary.....1

ESAP Review Comments.....2

I. Stormwater.....3

A. Stormwater Bioassays.....5
Fathead Minnow.....6
Ceriodaphnia.....6
Selenastrum.....6
Summary.....7

B. Stormwater Chemistry.....7
Summary.....7

II. Bay Water.....8

A. Bay Water Bioassays.....9
Inland Silverside.....9
Purple Sea Urchin.....10
Skeletonema.....10
Summary.....10

B. Bay Water Chemistry.....11
Summary.....11

III. Sediment.....11

A. Sediment Bioassays--Elutriate.....12
Mysid Shrimp.....13
Sanddab.....13
Bivalve (Pacific Oyster).....14
Summary.....14

B. Sediment Bioassays--Solid Phase.....15
Mysid Shrimp.....15
Amphipod.....15
Polychaete.....16
Summary.....16

C. Sediment Chemistry.....16
Reference Stations.....17
Sampling Stations--Surface Sediments....18
Sampling Stations-Subsurface Sediments..18

IV. Mussel Tissue.....19

A. Summary.....19

V. Conclusions.....20

VI. Recommendations.....21

VII. Attachments.....25

A. Memorandum from ORD regarding statistical analysis of bioassay data

B. Relationship of Sample Stations with Other Site Features (created from similar ESAP tables)

C. Summary Data Tables

1. Summary of Stormwater Bioassay and Chemistry Results
2. Summary of Bay Water Bioassay and Chemistry Results
3. Summary of Sediment Bioassay and Chemistry Results
4. Summary of Mussel Bioassay and Chemistry Results

EXECUTIVE SUMMARY

The Navy performed some sampling and analysis under the direction of the Environmental Sampling and Analysis Plan (ESAP) which was finalized on July 31, 1991. The study generally sought to evaluate the level of contamination present in the near shore environment of San Francisco Bay within the facility boundaries of Hunters Point Annex (HPA) in South San Francisco. The achievement of this objective was pursued through the use of both biological (laboratory bioassays using aquatic and benthic organisms) and chemical (laboratory chemical analysis) methods to analyze the potential contamination of surface water, sediments, and animal tissue. The results of the ESAP were submitted to the agencies on April 21, 1993.

In short, the bioassays performed for the ESAP were generally unsuccessful. Due to poor control survival, odd dose-response relationships, and other interfering factors, the data results were difficult to interpret and in many cases simply invalid. The study effort was so extensive, however, that rather than reject all of the results, the Environmental Protection Agency (EPA) suggests that in this case the data be reviewed more creatively than is generally recommended as a means of identifying broad areas of concern and establishing priorities for more focussed future studies. For example, the bioassay results suggested that at least two of the stormwater outfalls discharges were toxic to aquatic organisms. These two outfalls and their associated drainage areas should be prioritized for further monitoring. In addition, most of the sediment stations exhibited toxicity in sample pore-water. Thus, soluble contaminants should be prioritized for additional study. The relative failure of the bulk sediment bioassays suggest that the whole sediment should be further screened for toxicity.

The chemical analyses, on the other hand, were generally successful. The ESAP submittal, however, was missing several items related to the chemical analyses, such as: reference area water quality data and the full scan of analytical results for organic constituents. But, based on the submittal, several things are apparent.

First, stormwater discharged from Hunters Point Annex (HPA) to San Francisco Bay contains elevated levels of contaminants. Based mostly on inorganic data, Bay water in the vicinity of the stormwater discharges, on the other hand, have detectable levels of contaminants only. The levels are generally not levels of concern. Of course, the reference area data and additional organic data must still be evaluated before drawing final conclusions.

Second, bay sediment is significantly contaminated both on the surface and to a depth of 2.5 feet. Clearly, additional

effort must be made to determine the full nature and extent of this contamination, at least to a degree sufficient to determine an appropriate remedy, should the Ecological Risk Assessment (ERA) and Human Health Risk Assessment confirm the need for remediation.

Third, most of the mussel tissue collected contains elevated levels of contaminants. These results suggest that though Bay water samples do not contain contaminants at levels of initial concern, the levels are sufficient to cause bioaccumulation in organisms.

ESAP REVIEW COMMENTS

The ESAP set out a method for determining whether or not contaminants released from HPA to San Francisco Bay have caused toxicity to aquatic and benthic biota. The results of the ESAP were submitted in pieces to the agencies beginning in January 1993 through April 1993. A meeting at the end of January 1993 was held to discuss the preliminary findings. A follow-up meeting has not yet been scheduled. The EPA strongly recommends that such a meeting be scheduled in the near future.

The ESAP identified three separate sampling and analytical efforts for the purpose of measuring near shore toxicity. First, the ESAP endeavored to analyze the toxicity and chemical constituency of stormwater released from HPA to San Francisco Bay. Freshwater bioassay test species were used to measure toxicity from these sampling stations. As an aside, the stormwater component of the ESAP only looked at aqueous flows in the storm drains--not at settled sediment. Further, it only looked at flows associated with storm events--not those associated with on-going discharges to the storm sewer system.

The stormwater component of the ESAP further endeavored to analyze the toxicity and chemical constituency of Bay water from sites in close proximity to the sampled stormwater outfalls. Estuarine bioassay test species were used to measure toxicity from these sampling stations; and, a reference station in San Pablo Bay was identified for comparison to the bay water sampling results.

Second, the ESAP endeavored to analyze the toxicity and chemical constituency of near shore sediments which have likely been impacted by activities on and around HPA. Both elutriate and solid-phase sediment bioassays were performed on samples collected from 17 stations around the HPA--that is, both the sediment pore-water and sediment grains themselves were each analyzed for toxicity. Surface sediment samples were analyzed for toxicity while both surface and subsurface samples (-2.5 feet) were analyzed for chemical contamination.

Third, the ESAP endeavored to analyze the bioaccumulative potential of releases from HPA through the use of in-bay mussel bioassays. Mussels were planted near the 17 sediment stations and collected for tissue analysis after several weeks of exposure.

The results of the ESAP were meant to feed into the first phase of the ERA (Phase 1A). However, due to poor success of most of the bioassays, much of the data which was generated as part of the ESAP will not be of quantitative use in the ERA. While most of the chemical data is of acceptable quality, the bioassay data will have to be more broadly interpreted.

The following review is an attempt to begin the interpretive process which will be more completely formulated by the Navy in consultation with the agencies during Phase 1A of the ERA. Further, the review is an attempt to identify those areas of concern which should be addressed in additional field work as part of Phase 1B of the ERA--or by some other means as discussed among the parties.

I. STORMWATER

The impact of activities at the HPA on surface water quality is of concern, particularly the impact through stormwater runoff and other water discharges. Regarding this concern, a study conducted in 1989-90 as part of the Navy's on-going Installation Restoration (IR) investigation endeavored to: identify representative stormwater sampling sites, determine the quality of sediments within the storm drains, and determine the quality of water within the storm drains before and after a storm. The study was conducted as a means of "[characterizing] the chemical quality of stormwater runoff that discharges into San Francisco Bay during rainfall events." [Water Quality Investigation of Stormwater Drainage, 1991 (WQI, 1991)].

The results of the study were the basis for the stormwater component of the ESAP. The study indicated that in the comparison of the quality of water flowing into the stormwater sewers with that of sediments accumulated in the stormwater sewers, the sediments were a much more significant source of contamination. As result of the findings, the Navy recommended that the nature and extent of the sediment contamination within the stormwater sewer system be determined and the sediments potentially remediated. Further, the Navy concluded that storm events contributed higher concentrations of metals to the bay than did pre-storm discharges. The Navy recommended that the stormwater results be compared with likely National Pollutant Discharge Elimination System (NPDES) permit limits to determine the relative level of stormwater discharge compliance.

Though stormwater sewer sediments were identified as significantly contaminated, the source of the sewer sediment contamination was not identified--nor did the Navy, as part of the study conclusions, recommend that the sources be identified. Without identifying the sources of sewer sediment contamination, however, any removal of sewer sediments will likely provide only temporary relief from storm-related Bay contamination. That is, contaminants appear to enter the stormwater sewer system from either the indirect discharge of contaminants entrained in stormwater flows, the direct discharge of contaminants from HPA activities, or some combination of both. Until unauthorized direct discharges to the stormwater sewer system are ceased and random surface sources are identified and isolated, contaminated sediments are likely to continue to settle within the stormwater system and be discharged to the Bay during storm events.

The surface water component of the ESAP addresses neither the contaminated sewer sediments nor the identification of those contaminant sources which might be relevant to the stormwater contamination. Instead, the surface water component of the ESAP endeavors to evaluate "the potential toxicity of storm water runoff from HPA." (ESAP, 1991) Had the ESAP resulted in assurance that stormwater flows--including entrained sediments--were not toxic to organisms, then perhaps an argument could have been made that the contaminated sewer sediments needed neither further study nor remediation. However, not only did the results of the ESAP indicate the presence of contaminants in stormwater which exceed water quality criteria for aquatic and human health, but that test organisms suffer toxic effects when exposed to HPA stormwater, as well. The Parcel A Site Investigation report asserts that raw sewage has been discharged through the stormwater sewers system, thereby suggesting at least one possible source of the identified contamination.

Other sources of contamination may be found through the association of the stormwater sampling stations with known or suspected areas of release on the HPA facility which are otherwise under investigation through the Navy's IR program. Of the ten stormwater drainage basins on HPA--known as areas A through J--four are represented by stormwater sampling stations. The sampling stations are labelled ST-1 through ST-4 and are the same as those identified as part of the WQI, 1991.

According to the WQI, 1991, ST-1 represents flows from across Area D (35 acres) and collects runoff draining sites IR-6 (Tank Farm) and IR-19 (Building 901). ST-2 represents flows from across Area H (33 acres) and collects runoff draining sites IR-8 (Building 503 PCB Spill Area), IR-9 (Pickling and Plate Yard), PA-33, PA-37, and PA-44. ST-3 represents flows from across Area A (200 Acres) and collects stormwater draining sites IR-2 (Bay Fill Area), IR-4 (Scrap Yard), IR-5 (Old Transformer Storage Yard), IR-8 (Building 503 PCB Spill Area), IR-12 (Disposal

Trench), IR-13 (Old Commissary Site), PA-41, and PA-56. ST-4 represents flows from across Area E (30 acres) and collects runoff draining sites PA-58, PA-28, and PA-29. The drainage area sampled by ST-4 is known to have experienced significant historical industrial discharges. (Please note that the description of Area A is different than that which is contained in the ESAP. This discrepancy should be clarified).

The results of the ESAP indicate that stormwater from Areas A and E exert a definite toxic effect on aquatic organisms while stormwater from Areas D and H exerts a less certain effect. All four areas contribute levels of contaminants which exceed water quality criteria for aquatic and human health. The bioassay results and chemical analytical results are discussed separately below.

A. STORMWATER BIOASSAYS

Toxicity tests were conducted on samples taken from within the stormwater system (ST-1 through ST-4) at HPA. Toxicity was analyzed using several freshwater organisms: fathead minnow (*Pimephales promelas*) for survival and growth; cladoceran (*Ceriodaphnia dubia*) for survival and reproduction; and green alga (*Selenastrum capricornutum*) for reproduction. Each test was conducted using three replicates for each of the five serial dilutions (1%, 3%, 10%, 30%, 100%). Filtered reservoir water was used as the diluent. For two of the three stormwater bioassays, only one set of controls was run to compare to all of the test results.

There are a few problems with the test design as it relates to interpretation of the results. The large gap in concentrations between the 30% and 100% test solutions is such that it is difficult to interpret data which suggests that a 30% dilution is not toxic while a 100% sample is toxic. One cannot discern the slope of the dose-response curve nor the concentration at which there are no effects. Ultimately, this problem will impact the selection of clean-up criteria, if criteria must be developed.

Further, with only three replicates, the power of the statistical analysis is significantly reduced, particularly for non-parametric data results. A minimum of four replicates is probably necessary to achieve more robust statistical strength.

And finally, in those cases where only one set of controls was run for comparison to all of the sample test results, the ability to correctly evaluate the results for each stormwater sample is significantly reduced. [See Appendix A for a discussion of the bioassay statistical analyses, developed by EPA's Office of Research and Development (ORD)].

Of the three tests, only the fathead minnow test resulted in statistically valid data. Despite this fact, however, it is EPA's position that the other tests provide some valuable information. For the purpose of a preliminary screening, the 100% stormwater results were compared to control results for an indication of toxicity. The individual bioassays are discussed below.

1. Fathead Minnow

The fathead minnow test was generally considered to be valid. The results suggest that stormwater collected from ST-1 and ST-2 exerted no toxic effects on the minnow. However, 100% stormwater from ST-3 exerted a slight acute toxic effect and that from ST-4 exerted a more clear acute toxic effect.

2. Ceriodaphnia

The cladoceran test was considered invalid, primarily due to the lack of an observable dose-response relationship. The Navy suggests that factors other than sample dose may have affected the test organism. Looking exclusively at the 100% treatments, it appears that ST-2 and ST-4 exhibit both an acute and chronic toxic effect.

The text which describes the test design indicates that three replicates were run while the appendix related to the statistical analysis suggests that as many as 10 replicates may have been conducted. This discrepancy should be clarified.

3. Selenastrum

The green alga test was considered inconclusive due to low initial cell densities and poor documentation of the procedures. In addition, the dose-response relationships were oddly formed. Further, there apparently was no control for any of the four test samples. Instead, the control results obtained for the reference toxicant test are shown for comparison. All of these are problems due to poor laboratory performance and should be corrected in future studies.

Nonetheless, looking exclusively at the 100% treatments, it appears that ST-1, ST-2 and ST-3 exhibit a chronic toxic effect. In the statistical analysis attachments, it appears that the 100% treatment for ST-4 resulted in an algal cell density of 23.5 million cells/mL rather than the 27.4 which is printed in the summary table. With this correction, it may be that even ST-4 exhibits a toxic effect when compared to the 26.1 million cells/mL observed in the reference toxicant control.

4. Summary (See Appendix C, Table 1)

a. The analysis of samples from ST-1 indicate an unconfirmed reduction in the reproduction of green algae in the presence of contaminated stormwater.

b. The analysis of samples from ST-2 indicate an unconfirmed reduction in survival and reproduction of the cladoceran in the presence of contaminated stormwater. Further, the analysis indicated an unconfirmed reduction in reproduction of green algae in the presence of contaminated stormwater.

c. The analysis of samples from ST-3 indicate a confirmed reduction in survival of fathead minnow in 100% stormwater runoff. Further, the analysis indicates an unconfirmed reduction in the reproduction of green algae.

d. The analysis of samples from ST-4 indicate a confirmed reduction in survival of fathead minnow in 100% stormwater runoff. Further, the analysis indicates an unconfirmed reduction in the survival and reproduction of the cladoceran. Unconfirmed toxicity may also be indicated in the 100% treatment using green algae, though the statistical significance of the result is unclear.

B. STORMWATER CHEMISTRY

Stormwater samples were collected from ST-1 through ST-4 for chemical analysis in addition to the bioassay testing which was described above. The EPA compared the results of the chemical analyses to EPA's published Water Quality Criteria (WQC). This comparison indicates the presence in HPA stormwater of contaminants in levels of concern.

Criteria have not been developed for all of the contaminants which were measured as part of the ESAP. Thus, the following observations do not illustrate the full breadth of contamination which may be present. Furthermore, judging by the data which was submitted, either the scan of organic contaminants measured as part of the ESAP was extremely limited, or the full scan of organic data was simply not submitted. Nonetheless, the study results indicate a need for further investigation regarding the stormwater pathway for the movement of contaminants from on shore sources to the near shore ecosystem.

1. Summary (See Appendix C, Table 1)

a. The analysis of samples from ST-1 indicate the presence of lead in levels which exceed the WQC for acute and chronic toxicity for freshwater organisms and the 10^{-6} risk level for cancer to humans through the ingestion of water and organisms. The analysis also indicates the presence of copper in

levels which exceed the WQC for acute and chronic toxicity to freshwater organisms and zinc which exceed the WQC for chronic toxicity to freshwater organisms. In addition, barium, manganese, zinc, monobutyltin, dibutyltin, tributyltin, fecal coliform bacteria were detected in the stormwater at this station.

b. The analysis of samples from ST-2 indicate the presence of lead in levels which exceed the WQC for acute and chronic toxicity for freshwater organisms and the 10^{-6} risk level for cancer to humans through the ingestion of water and organisms. The analysis also indicates the presence of copper in levels which exceed the WQC for acute and chronic toxicity to freshwater organisms and zinc in levels which exceed the WQC for acute and chronic toxicity to freshwater organisms. In addition, aluminum, barium, chromium, manganese, monobutyltin, dibutyltin, tributyltin, and fecal coliform bacteria were detected in the stormwater at this station.

c. The analysis of samples from ST-3 indicate the presence of lead in levels which exceed the WQC for chronic toxicity to freshwater organisms. The analysis also indicates the presence of copper in levels which exceed the WQC for chronic and acute toxicity to freshwater organisms; manganese in levels which exceed the 10^{-6} risk level of cancer to humans through the ingestion of both water and organisms and through the ingestion of organisms alone; and zinc in levels which exceed the WQC for chronic and acute toxicity to freshwater organisms. In addition, aluminum, barium, vanadium, monobutyltin, dibutyltin, tributyltin, and fecal coliform bacteria was detected in the stormwater at this station.

d. The analysis of samples from ST-4 indicate the presence of lead in levels which exceed the WQC for chronic toxicity to freshwater organisms. The analysis also indicates the presence of copper in levels which exceed the WQC for acute and chronic toxicity to freshwater organisms and zinc in levels which exceed the WQC for acute and chronic toxicity to freshwater organisms. In addition, aluminum, barium, manganese, monobutyltin, dibutyltin, tributyltin, trichloroethene and fecal coliform bacteria were detected in the stormwater at this station.

II. BAY WATER

A second method by which the ESAP endeavored to analyze the potential toxicity of storm water runoff from HPA was to conduct toxicity tests on Bay water samples taken in the vicinity of those stormwater outfalls associated with the stormwater sampling effort. The Bay water samples were taken from four sampling stations identified as B-1 through B-4. B-1 was located in the vicinity of ST-1, B-2 in the vicinity of ST-2, etc.

Only one of the Bay water sampling stations exhibited any clear toxicity. An inconclusive test suggested the occurrence of chronic toxicity at three of the sampling stations. And, contaminants were detected in the chemical analyses in all of the Bay water samples; but, none of them exceeded WQC for marine organisms.

While the ESAP identified a reference sampling location in San Pablo Bay to compare to Bay water toxicity and chemical constituency, no such results appear in the materials thus far submitted. It is unclear whether reference samples were collected, but the results not submitted, or no such samples were ever collected. This discrepancy should be clarified.

A. BAY WATER BIOASSAYS

Toxicity in Bay water was measured using estuarine bioassay species, including: the inland silverside (*Menidia beryllina*) for survival and growth; the purple sea urchin (*Strongylocentrotus purpuratus*) for fertilization; and a marine diatom (*Skeletonema costatum*) for reproduction. Each test was conducted using three replicates for each of five dilutions in a logarithmic series (1%, 3%, 10%, 30%, 100%). Synthetic seawater was used as diluent and only one set of controls was run for all of the samples and dilutions.

The problems associated with the test design are the same as those mentioned for the stormwater bioassays. In addition, however, the manufactured seawater diluent used in the Bay water bioassays may have exerted its own toxic effect which acted to complicate the test results. In many instances, toxicity appeared to decrease as the relative concentration of effluent increased, suggesting such an interference. The use of natural seawater--or reconstituted natural seawater--in future such tests is highly recommended.

1. Inland Silverside

The inland silverside test was considered valid. However, there was only one set of controls run for a total of 60 test chambers. According to EPA's ORD (see Appendix A), one set of controls for each discrete sample is recommended--one control per 15 treatments and replicates.

The 100% sample from B-3 exhibited 83% survival as compared to 100% survival in the control. The report indicated that only the 30% sample take from B-2 exhibited a result which was significantly different than the control.

2. Purple Sea Urchin

The purple sea urchin test was also considered valid, though the test acceptability criterion for control survival was apparently not met. In this test also, only one set of three control replicates was run along side 60 test chambers when one set of control replicates for each set of sample treatments and replicates might have been advisable.

The 100% samples exhibited toxicity in all four samples though a reasonable dose response curve was only discernable in the B-2 test.

3. Skeletonema

The diatom test was considered inconclusive because the diluent may have inhibited diatom growth and the bioassay laboratory did not report true control data, only performance control data. In this test also, only one set of controls was run along side 60 test chambers when control replicates for each set of sample treatments and replicates would have been advisable.

The 100% samples for B-1, B-2 and B-3 exhibited reduced algal growth. Algal growth in the B-4 samples seemed to increase with sample strength.

4. Summary (See Appendix C, Table 2)

a. The analysis of samples from B-1 indicate a reduction in fertilization of sea urchin ova in 100% stormwater runoff. The analysis, however, results in only a vague dose-response relationship. Diatom growth appears reduced in the 100% treatment, though no dose-response relationship is demonstrated.

b. The analysis of samples from B-2 indicate a reduction in fertilization of sea urchin ova in 100% stormwater runoff. The analysis, however, results in only a vague dose-response relationship. Diatom growth appears reduced in the 100% treatment, through no dose-response relationship is demonstrated.

c. The analysis of samples from B-3 indicate a reduction in survival of inland silverside in 100% stormwater runoff, though the significance of the results is not clear. The sea urchin test indicated a reduction in fertilization in 100% stormwater runoff, however, only a vague dose-response relationship is observed. Further, the analysis indicates a reduction in algae growth, but neither is a dose-response relationship demonstrated nor is the significance of the reduced growth clear.

d. The analysis of samples from B-4 indicate a reduction in fertilization of sea urchin ova in 100% stormwater runoff. The analysis, however, provides no dose-response relationship.

B. BAY WATER CHEMISTRY

The results of the Bay water samples were meant to indicate the effect on water column concentrations of contaminants as contaminated stormwater is released to and mixes with Bay water.

As a general comment related to the chemical data summary tables provided as part of the ESAP data submittal, the stormwater and Bay water results were inappropriately combined in the calculation of minimum, maximum and mean values. As distinct surface water types--freshwater and brackish water--a summary table for each, separately, should have been developed.

1. Summary (See Appendix C, Table 2)

a. The analysis of samples from B-1 indicate the detection of barium, chromium, manganese, zinc, monobutyltin, dibutyltin, and tributyltin. The lead detection limit was above the WQC for that compound. Otherwise, no compounds were detected at levels which exceed WQC.

b. The analysis of samples from B-2 indicate the detection of barium, manganese, zinc, monobutyltin, dibutyltin, and tributyltin. The lead detection limit was above the WQC for that compound. Otherwise, no compounds were detected at levels which exceed WQC.

c. The analysis of samples from B-3 indicate the detection of aluminum, barium, manganese, zinc, dibutyltin and tributyltin. The lead detection limit was above the WQC for that compound. Otherwise, no compounds were detected at levels which exceed WQC.

d. The analysis of samples from B-4 indicate the detection of barium, zinc, monobutyltin, dibutyltin, and tributyltin. The lead detection limit was above the WQC for that compound. Otherwise, no compounds were detected at levels which exceed WQC.

III. SEDIMENT

The impact of HPA activities on near shore sediments is of concern. Regarding this concern, the Navy included as part of its investigations of Operable Unit I (OU I) and OU IV, near shore sediment sampling and chemical analysis. The results were published in a Technical Memorandum dated January 29, 1993 and

indicated high level of inorganic and organic contaminants.

The ESAP broadened the scope of OU I and OU IV investigation to look at both biological toxicity and chemical contamination of sediment stations all along the edges of HPA, not just on the shores of OU I and OU IV.

The ESAP identified two methods of measuring toxicity of near shore sediments: elutriate and solid phase bioassay procedures. The elutriate bioassays were meant to identify toxicity which might be associated with sediment pore water. And, the solid-phase bioassays were meant to identify toxicity which might be associated with the sediment particles, themselves. Further, the ESAP identified both surface and subsurface sediment contamination a concern. Sediment samples taken on the surface and at 2.5 feet below the sediment surface were therefore chemically analyzed.

The ESAP established 17 sediment sampling stations around the perimeter of the HPA in the intertidal zone: S-1 through S-17. Appendix __ illustrates the relationship among the stormwater, Bay water, sediment and mussel sampling stations, as well as with other features of the HPA site. The sediment stations appear to be well placed around the facility's perimeter so as to identify sediment contamination due to either stormwater discharges or more random discharges to the bay from the HPA facility, including from the dry docks.

In addition to the 17 sampling stations, the ESAP also established three sediment reference stations: just south of Candlestick Park in San Francisco Bay (ACED-1), just north of Sierra Point in San Francisco Bay (ACED-2) and at the four-county junction in the middle of San Pablo Bay (ACED-3). There is, unfortunately, no indication that these reference areas were pre-screened to determine the appropriateness of their use as references. A preliminary study of the toxicity and chemical constituency of these stations would have been useful in making a more informed decision regarding their appropriateness.

Sediment samples were collected during the course of several field days. Ten random surface grab samples were collected from each sediment sampling station and composited in the laboratory. One sediment core sample was collected from each sediment sampling station and a discrete portion from between 30" and 36" was extracted for analysis. The surface sediment samples were analyzed both by biological and chemical methods. The subsurface samples were analyzed by chemical methods, only.

A. SEDIMENT BIOASSAYS--ELUTRIATE

The toxicity of sediment-pore water was analyzed using several estuarine organisms, including: mysid shrimp

(*Holmesimysis costata*) for survival; sanddab (*Citharicthys stigmaeus*) for survival; and a bivalve (*Crassostrea gigas*) for embryo survival and normality. Based on the data validation process, only the bivalve test was considered to be valid. However, as part of EPA's broader screening method, those tests which exhibited toxicity in the 100% test are noted.

Difference than the stormwater and Bay water bioassays, each test was conducted using five replicates for each of the three serial dilutions (10%, 50%, and 100%). Natural seawater and sediment from Bodega Bay was used in the preparation of the samples and control. Tests were conducted in two groups: one including the three reference stations and the other not. One set of controls were run for each group.

There are a couple of problems with the test design as it relates to interpretation of the results. Without pre-screening of Bodega Bay water and sediments for toxicity and chemical constituency, it is unclear that such water and sediments can be declared clean for the purpose of test dilution and control. Some preliminary scoping would have been useful for identifying truly clean water and sediments. Further, splitting the test samples into two groups disallows a very accurate comparison of one site to another. In particular, in those cases where groups of sediment stations share a potential upgradient contaminant source, it would have been useful to have those samples all tested within the same group. Further, since the reference station samples were tested in Group 1 only, Group 2 tests have no relative toxicity standard for comparison.

1. Mysid Shrimp

The mysid shrimp test was generally considered to be invalid because of low sediment control survival in both of the groups. This may have been related to the quality of Bodega Bay sediments which was apparently not analyzed prior to its use in this test. No statistical analysis was conducted on the results because of the low control survivability. However, despite the poor control survivability, many of the test samples resulted in a reasonable dose-response curve with apparent reduction in survival concurrent with increases in sample concentration. Those stations which exhibit such a relationship and appear to exhibit a toxicity which is greater than the control are as follows: IGS-2, FS-3, S-1, S-2, S-3, S-5, S-6, S-7, S-9, S-12, S-14, S-15, S-16, and S-17. Both reference and sample stations exhibit a similar range of toxic response.

2. Sanddab

The sanddab test was generally considered to be inconclusive because of low control survival. A statistical analysis was nonetheless conducted on the test results because

the low control survival was considered to be marginal. Despite the poor control survival, many of the test samples resulted in a reasonable dose-response curve with apparent reduction in survival concurrent with increases in sample concentration. Those stations which exhibit such a relationship and appear to exhibit a toxicity which is greater than the control are as follows: S-4, S-6, S-8, S-9, S-10, S-12, and S-16.

3. Bivalve (Pacific Oyster)

The oyster test was generally considered to be valid. However, the narrative description of the test results indicate that the acceptance criterion for normal shape was met for Group 1 tests but not Group 2 tests. Yet, the summary indicates that the Group 1 control had a rate of larval abnormality of 17.4%--greater than the 10% allowed. The summary for Group 2 indicates that the control had a rate of larval abnormality of 13.9%--also greater than the 10% allowed, though still less than Group 1. This discrepancy should be clarified.

Despite the fact that the test acceptability criterion for larval normality was not met, the dose-response patterns of the chronic toxicity test appear to be relatively normal. The following stations exhibit a reasonable dose-response curve and rates of abnormality which are greater than the control: -1, -2, -3, S-1, S-2, S-3, S-4, S-8, S-10, S-13, S-14, S-15, and S-16. Of those, S-1, S-3, S-15, and S-16 show greater rates of abnormality than the reference area stations.

There are also several stations which exhibit both a reasonable dose-response and level of mortality which is greater than the control. Those stations are: S-1, S-2, S-4, S-13, and S-15. The reference areas exhibited a rate of mortality similar to that of the control.

4. Summary (See Appendix C, Table 3)

a. The following stations may have been acutely toxic to mysid shrimp: S-1, S-2, S-3, S-5, S-6, S-7, S-9, S-12, S-14, S-15, S-16 and S-17. Because the test was generally considered to be invalid, these observations are unsubstantiated.

b. The following stations may have been acutely toxic to the sanddab: S-4, S-6, S-8, S- 9, S-10, S-12, and S-16. Because the test was generally considered to be inconclusive, these observations are unsubstantiated.

c. The following stations may have been acutely toxic to the bivalve: S-1, S-2, S-5, S-10, S-13, S-15 and S-16. The test was considered to be valid, but these conclusions were not drawn. As such these observations are unsubstantiated.

d. The following stations were chronically toxic to the bivalve: S-1, S-3, S-5, S-6, S-8, S-10, S-12, S-13, S-14, S-15, and S-16.

e. In short, the pore water of sediment samples from every sediment station has levels of contaminants which may cause toxicity to aquatic organisms.

B. SEDIMENT BIOASSAYS--SOLID PHASE

The solid phase sediment bioassays were tested on 100% sample concentrations in five replicates. Sediment from Bodega Bay was used for a control. And, as with the other sediment bioassays, two groups were run--one with the reference samples and seven test samples and the other with the remaining test samples.

As with the elutriate test, there is no evidence that Bodega Bay is a reasonable source for control sediments. Unlike the elutriate test, however, each sample test is only run as a 100% test and thus suffers no interference from diluent contamination. Though the solid phase sediment bioassays offer an indication of the rate of survival of organisms in bay sediment, without concurrent reference station testing in each group, only Group 1 results can be formally compared to test sample results.

1. Mysid Shrimp

The mysid shrimp bioassay was generally considered to be valid. Due to inadvertent turbulence from aeration, however, reviewers were instructed to view the data cautiously because the results suggest unforeseen test conditions. In particular, the difference between Group 1 results and Groups 2 results suggests that Group 1 tests suffered some interference that Group 2 tests did not. Group 1 results, at face value, suggest that all of the reference stations and S-1, S-3, S-6, S-8, and S-14 through S-16 exhibit significant mortality in comparison to the control. All Group 2 tests, on the other hand, suggest that S-2, S-4, S-5, S-7, S-9 through S-13, and S-17 suffer no significant reduced survivability.

2. Amphipod

The amphipod bioassay was generally considered to be valid, though the results are identified as inconclusive due to the large rate of mortality suffered on the last day of the test. At face value, all reference and test stations, except S-10 perhaps, exhibit significant reduced survivability.

3. Polychaete

The polychaete bioassay was generally considered to be valid, but the results were inconclusive because of the large rate of mortality suffered on the last day of the test. At face value, all reference and test stations exhibited significant reduced survivability.

4. Summary (See Appendix C, Table 3)

a. None of the solid phase sediment bioassays were fully successful. There is no obvious method for using the available data to draw even general conclusions. The failures of these tests appear to be related to lab implementation of the bioassay protocols.

C. SEDIMENT CHEMISTRY

Sediment chemistry was measured for both surface and subsurface sediments. Surface sediments were analyzed as a composite of ten randomly collected samples which should be representative of the entire sediment sampling station. In addition, one subsurface sample was analyzed. No map indicating the location of subsurface sampling was provided in the data submittal.

As part of EPA's screening of the sediment data, the results of the sediment sampling effort were compared to the National Oceanic and Atmospheric Administration's (NOAA) biological effect range values as published in NOAA Technical Memorandum NOS OMA 52, August 1991. The low and median values associated with the effects range indicate the lower 10 percentile concentration at which biological effects occur and the median concentration at which biological effects occur, respectively. They are abbreviated ER-L for effects range-low and ER-M for effect range-median. The concentration numbers are useful as a screening mechanism to identify those sampling sites at which there is clear cause for concern. They are used here for that purpose, only.

As with EPA's WQC, it is important to note that NOAA has not developed ER-Ls and ER-Ms for all of the contaminants which were measured as part of the ESAP. Thus, they can not be used as the sole method of site screening. In this review, comparison of the chemical data to ER-Ls and ER-Ms has identified sites of concern and the need for improved detection limits in future investigatory work.

The results were also compared to the arithmetic mean of the reference area values for each contaminant. Non-detects were included in the mean as zero, no standard deviation was calculated, and no formal statistical analysis was performed.

Thus, this is a preliminary screening level effort, only.

Related to the interpretation of the summary tables submitted, it is unclear how non-detect data was incorporated into the minimum, maximum and mean values calculated for these tables. For example, are they included in the mean as zero or as the Method Detection Limit? There also is no similar analysis conducted on the reference area data such that comparisons of the reference area data to the sampling station data can be conducted. This matter should be clarified.

1. Reference Stations

Data was collected from three reference areas: just south of Candlestick Park in San Francisco Bay (-1), just north of Sierra Point in San Francisco Bay (-2) and at the four-county junction in the middle of San Pablo Bay (-3). Two sets of data were collected from -2 such that four data points are available for reference. A mentioned above, there was no initial screening to determine the appropriateness of these stations for reference.

In general, the data from these sites indicates that -2 has higher levels of inorganic contaminants than do -1 or -3. All stations have levels of lead and nickel which exceed NOAA's ER-Ms and levels of mercury which exceed the ER-L. Station -2 has levels of chromium which exceed NOAA's ER-L and -2 and -3 have levels of zinc which exceed the ER-L. Of interest is that several contaminants are found at higher concentrations at depth than on the surface, suggesting historically greater contaminant loading.

It is unclear whether or not these levels of contaminants can be considered "ambient" levels in the San Francisco/San Pablo Bay area--or if the chosen reference sites are themselves subject to some source of contamination which is not otherwise in effect at the Hunter's Point Annex.

The levels of lead, in particular, are an order of magnitude higher than those found at either HPA or in the Regional Water Quality Control Board's ambient monitoring program as published in December 1992 as "San Francisco Bay Pilot Regional Monitoring Program 1991-1992: Summary Progress Report." It may be that the levels reported in the data are in fact a typographical error requiring a decimal point which is presently not included. The nickel and zinc concentrations generally follow the trends identified by the Regional Board. The Regional Board did not track mercury concentration trends.

2. Sampling Stations--Surface Sediments

Surface sediments were collected as 10 separate grab samples across the sampling station surface all of which were later composited in the laboratory for analysis.

As a general matter, a comparison of the reference area data to the sampling data strongly suggests that activities at HPA have impacted the quality of San Francisco Bay sediments beyond ambient contamination.

3. Sampling Stations--Subsurface Sediments

Subsurface sediments were collected as a single core sample from each sediment station from which a discrete sample from between 30" to 36" was collected for chemical analysis. The data package does not indicate where within the sampling station each core was taken. A map indicating those locations should be submitted.

As a general matter, for many contaminants, it appears that historical contamination, a measured in the subsurface samples, was greater than current contamination, a measured in surface samples. The comparison of surface composite samples and subsurface grab samples, however, complicates the matter.

4. Summary (See Appendix C, Table 3)

The volume of data generated as part of the sediment chemistry study makes discrete summary remarks difficult. However, the following are general observations.

a. Nickel and pyrene are found at levels which exceed ER-Ms at nearly every sediment station.

b. Mercury is found at levels which exceed ER-Ls at nearly every sediment station.

c. Dibutyltin and Tributyltin are found at levels which exceed the mean reference are value at nearly every sediment station.

d. With few exceptions, contaminants found in surface samples were generally found in subsurface samples, as well.

e. The stations with 15 or more contaminants at elevated levels include: S-4, S-10, S-11, S-12, S-13, and S-14. As a preliminary screening, these can be considered the areas of most significant contamination.

f. The stations with 5 to 14 contaminants at elevated levels include: S-1, S-2, S-5, S-6, S-8, S-9, S-15, S-16, and S-17. As a preliminary screening, these can be considered the areas of moderate contamination.

g. The stations with less than 5 contaminants at elevated levels include: S-3 and S-7. As a preliminary screening, these can be considered the areas of limited contamination.

IV. MUSSEL TISSUE

The impact of HPA activities on biological health is of concern. The bioaccumulation-component of the ESAP was an attempt to determine whether such an impact was occurring. In particular, the ESAP proposed to evaluate "the potential presence of [metals, SOCs, organochlorine pesticides and PCBs, and tributyltin] in the San Francisco Bay surrounding HPA, and their potential for bioaccumulation into aquatic organisms...by measuring the chemical uptake of these substances into the mussel, *Mytilus californianus*." (ESAP, 1991)

Bioaccumulation was measured in two separate mussel deployments (4/7/92 and 9/9/92) by placing bags of 50 mussels at 17 stations which corresponded with each of the 17 sediment sampling stations. Mussels were also placed at two reference areas corresponding to RS-1 and RS-2 used in the sediment quality study. The reference station in San Pablo Bay was not used in the bioaccumulation study.

Fifteen individuals were composited for metals analysis, 20 for organic analysis, 5 for radioactivity, and 10 extra were included in case of mortality. Mussels were collected from Bodega Head. The results from the reference areas are presented, but neither the radiological data, much of the planned organic data, nor the background data is contained in the analytical results table. This information should be forwarded to the agencies.

As part of EPA's data screening effort, the results of the mussel study were compared to the arithmetic mean of the reference area results. The Phase 1A ERA presentation should include a formal statistical analysis of these results.

A. Summary (See Appendix C, Table 4)

1. Many chemicals were found in the tissue of mussels deployed at the reference stations. The tissue of mussels deployed at RS-1 (Candlestick Park) exhibited levels of contaminants which were higher than those found in mussel tissue at RS-2 and thus for most contaminants offer the upper range of

reference values. Non-detected values are included in the mean as a zero. A standard deviation was not calculated.

2. As a general comment, the data collected beginning on 4/7/92 does not fully correspond to that which was collected beginning on 9/9/92. Neither testing period consistently exhibited more bioaccumulation than the other.

3. All of the sampling stations exhibited levels of tissue accumulation of inorganic contaminants which were in excess of those identified by the mean of the reference area values.

4. Several of the stations exhibit levels of tissue accumulation of organic contaminants which were in excess of those identified by the mean of the reference area values. Those stations are: M-1, M-3, M-9, M-10, M-11, M-12, M-13, M-14, and M-15.

5. Several of the stations were analyzed for diethyl phthalate at a detection limit which was higher than the mean reference area value and thus could not be fully evaluated for that compound. Those stations are: M-2, M-3, M-5, M-6, and M-17.

6. With a total of 29 compounds which were detected at any station, the following stations exhibited high level of 20-29 of them, as a mean of the two collection period results: M-1, M-2, and M-4. As a preliminary screening, these can be considered the areas of most significant bioaccumulation.

7. With a total of 29 compound which were detected at any station, the following stations exhibited high levels of 10-19 of them, as a mean of the two collection period results: M-3, M-5 through M-12, and M-14 through M-17. As a preliminary screening, these can be considered the areas of moderate bioaccumulation.

8. With a total of 29 compounds which were detected at any station, only M-13 exhibited high levels of 0-9 of them, as a mean of the two collection period results. As a preliminary screening, this station can be considered an area of limited bioaccumulation.

V. CONCLUSIONS:

The results of the stormwater component of the ESAP provide evidence that contaminants are entering San Francisco Bay from HPA through the stormwater drainage system. Both the biological and chemical analyses support this claim, though the biological analyses provide less certainty. Of the four stations, Station ST-4 appears to represent the worst of the stormwater conditions: measured acute toxicity and elevated metals and TCE. The ESAP

stormwater data should be used to prioritize future efforts.

The results of the stormwater component further suggest that the discharge of stormwater to San Francisco Bay has only minimal effects on the Bay water quality itself. The data presented regarding bay water quality, however, is relatively limited. The water chemistry, in particular, is not complete.

The results of the sediment component of the ESAP suggests that considerable contamination of bay sediments has occurred, both historically and presently. All of the stations have levels of contaminants which exceed reference area levels. Further, most stations have levels which exceed biological effect levels, both at depth and on the surface. Thus, it appears that the near shore sediments must be folded into the rest of the on-going CERCLA activities at HPA with the goal of sediment remediation in mind. The method for best accomplishing such a goal should the subject of a meeting in the near future.

Associated with the chemical data is corroborative biological data; though, admittedly the data is of questionable validity. Nonetheless, it suggests that not only does the contamination exceed that which is found at reference areas, and that which is expected to cause biological effects, but in fact actually causes biological effects.

Finally, the bioaccumulation component of the ESAP suggests that contaminant releases in and around all of the mussel stations cause bioaccumulation of those contaminants in mussel tissue. Thus, the evidence suggests that despite the preliminary findings of the bay water quality component of the ESAP, levels of contaminants are present in the water column which though perhaps not lethal, do in fact bioaccumulate.

VI. RECOMMENDATIONS

1. Identify the source of contaminated sediments observed as part of the WQI, 1991.

2. Identify, characterize, and cease all unauthorized discharges to the stormwater sewer system.

3. Future bioassays should include a control for each definitive test, where possible.

4. Future bioassays should include an appropriate number of replicates to ensure the ability to use non-parametric statistical methods, if such methods are warranted. The number of dilutions should be balanced with the number of replicates so that robust statistical methods are available for use, a dose-response relationship can be observed, and clean up criteria can be developed, where needed.

5. The ESAP proposed to only use a laboratory for conducting the bioassays which has been approved by the RWQCB as a bioassay laboratory for chronic toxicity testing and has participated in the EPA "Round Robin" testing program with acceptable results. It should be noted that Aqua Terra, who performed the ESAP bioassays, did not participate in the EPA "Round-Robin" testing program with acceptable results. They inappropriately applied statistical methods which resulted in their failure of the "Round Robin" test. Future bioassays should be conducted by a laboratory which is held to the standard set by the ESAP.

6. The bioassay data generated by the ESAP should be evaluated broadly--and creatively--to ensure that all available value is obtained from the extensive effort. In particular, the bioassay data can be used to identify areas of concern to focus future studies.

7. The invalid or inconclusive bioassays should be reviewed to determine the source of their relative failures. A duplication of these failures should then be avoided in future bioassay performance.

8. The discrepancy between the summary table and the statistical analysis documentation for the cladoceran test should be clarified.

9. In the future, if initial bioassay test conditions are not acceptable (e.g., low alga cell density), then early corrections should be made so that the entire test is not wasted.

10. The summary table depicting the green algae results should be revised to correct the ST-4, 100% treatment result.

11. All of the organic data for stormwater, Bay water, sediment and mussel analyses should be submitted to agencies, if it has not X been submitted. In the event that all the data has been submitted, an explanation for the availability of only limited organic data should be provided.

12. The Bay water reference data should be submitted to the agencies.

13. Future estuarine or marine bioassays should use natural or reconstituted natural seawater instead of manufactured seawater. It appears that manufactured seawater caused toxic interference in these tests.

14. Future Bay water chemical analyses should be conducted at detection limits low enough to measure exceedences of screening criteria, such as EPA's WQC and the States WQC.

15. Statistical summaries of chemical data should not lump

together different sample sources. For example, freshwater stormwater samples should be summarized separately from brackish Bay water samples. Further, the statistical summaries should include the median and geometric mean, as well as the arithmetic mean.

16. An assessment of how each of the sampling stations are related to one another and to flows from HPA should be compiled as part of Phase 1 of the ERA. The discrepancies between the WQI, 1991 and ESAP, 1991 regarding this matter should be clarified.

17. Prior to initiating future studies, reference areas should be screened to determine their appropriateness as references.

18. Greater care should be taken in the separation of samples into groups for analysis. For example, sample stations which might generally reflect the activities of a shared source should be grouped together so that the information gathered from each individual test can be combined together. Further, reference area samples should be tested in each group so that the reference data can in fact be used to compare to all sample test results.

19. A map depicting the location within each sediment sample area from which sediment cores were taken should be submitted to the agencies.

20. Prior to initiating future studies, natural diluent sources--whether it be water or sediment--should be screened to determine that the material is in fact devoid of interfering toxins.

21. The re-evaluation of the oyster larvae acceptance criteria should be made.

22. Future sediment bioassays should be implemented by a laboratory with a demonstrated ability to implement and interpret them properly.

23. All statistical summary tables should include a note which describes how non-detect data is included in the summary.

24. A statistical summary of reference area sample results should be included in future data packages--separate from the test sample results.

25. The lead levels identified in reference sediment samples may require a decimal point which not currently present. This data should be reviewed for accuracy.

26. All of the ESAP data should be more formally correlated as part of the Phase 1 ERA. In particular, a correlation of on-shore sources with near shore contamination may provide some

indication of where opportunities for source control exist.

27. The results of the mussel radiation screening should be submitted to the agencies.

28. A meeting should be scheduled to talk both about the substance of the ESAP results and its application to future investigatory and remedial efforts.

Appendix A

ORD's Review of the Navy's Bioassay Statistical Analysis

TECHNOLOGY APPLICATIONS, INC.

3411 Church Street
Cincinnati, Ohio 45244
(513) 533-8114
Fax: (513) 533-8181

DATE: July 26, 1993

TO: James M. Lazorchak, EMSL-Cincinnati

FROM: Laura C. Gast, TAI/EMSL-Cincinnati Statistical Support LCG

SUBJECT: Comments on "Data Validation Summary Analysis, Naval Station Treasure Island, Hunter's Point Annex, San Francisco, California"

This provides review comments on the statistical data analysis contained in the above document. Overall, the statistical analyses performed are the standard analysis techniques recommended for aquatic toxicity test data in the U. S. EPA acute and chronic toxicity testing methods manuals. One exception is a failure to perform checks of the normality and homogeneity of variance assumptions for the two-sample t tests used to analyze the sediment elutriate and sediment test data. This and several additional concerns regarding the report's statistical analysis are listed below.

COMMENTS:

* In each of the stormwater bioassays except the cladoceran tests, a single set of control replicates is used for comparison with all the stormwater dilution series results. This raises two concerns. First, three control replicates is an extremely small number to ensure quality results when being compared to four separate stormwater samples, with three replicates for each of their five dilutions. While it seems to be common practice to use a single set of control replicates when performing several effluent profile tests simultaneously, I am not aware of other instances where a set of definitive tests shares a single, common, set of control replicates. It is also unclear whether the tests on the four samples were performed at the same time, or if they were run at different times, with the controls present on only one occasion.

  The second concern is test design and randomization. It is standard practice to randomly arrange all the replicates for a definitive toxicity test on a test tray or board. It is unclear, with the single set of control group replicates, how the test chambers were randomized, if at all. With four stormwater samples and just three control replicates, there could not have been even one control chamber per stormwater sample test board.

* It is unfortunate that only three replicates per sample concentration were used in the stormwater bioassays. In such a situation, nonparametric analysis techniques cannot be used, necessitating the use of Dunnett's Procedure even in cases where one or both of the normality and homogeneity of variance assumptions are violated. Results must be viewed with extreme

caution in these cases.

* Many of the bioassays reported in this document contain results with unclear dose-response relationships where high concentrations of a stormwater sample did not show toxicity while intermediate dilutions demonstrated significant toxicity, complicating data analysis. It is standard procedure to remove concentrations with significant mortality versus the control to be removed from further analysis of effects such as growth and reproduction. This practice is counter-intuitive when the significant concentrations are intermediate concentrations rather than high concentrations. It is then best to continue the analysis with response data from all concentrations except significantly toxic high concentrations as in this report.

* The sediment elutriate and sediment test data were analyzed with individual two-sample t tests. In this situation the Type I (false positive) error rate is controlled for each comparison, but not for the set of comparisons as a whole. Therefore, statements of statistical significance of comparisons may only be made on an individual sample basis, not for the entire set of samples, because the overall Type I error rate was not controlled. This is a subtle point, but bears mentioning.

  In a worst case scenerio, which assumes each comparison is independent, the upper bound of the probability of at least one Type I error out of the 17 comparisons is: $1 - (1 - 0.05)^{17} = 0.582$. While the actual probability could be considerably less than 58.2%, it is still a sizeable chance of declaring at least one site to be "toxic" when, in reality, it is not. Further, since these tests were performed in two test groups, presumably because the samples were collected at different times, multiple comparisons could be performed only on the samples within each group, not on the overall collection of samples.

  In addition, these two-sample t tests carry assumptions of normality and homogeneity of variance similar to Dunnett's Procedure. It appears that these assumptions were not tested prior to performing the t tests. The toxicity testing manuals provide an unequal variance case modified t test as well as a nonparametric alternative, Wilcoxon's rank sum test. These alternatives are preferrable to the use of Dunnett's Procedure in Section 11, the Mysid sediment tests, for individual comparisons since these tests were performed with five replicates per sample.

* The statistical analysis sections concerning the amphipod and polychaete sediment tests indicate that since organism survival was signficantly lower than the respective control in each sample (except one amphipod test site) that the results "indicate that apparent toxicity at HPA is no greater than toxicity at the reference sites." This statement is based on inspection rather than on formal tests of site responses versus reference site responses. Further, since all three reference sites were part of the group one samples, they cannot be statistically compared to the group two samples, since test conditions for the two groups could have been different and thus not directly comparable.

* In the sediment elutriate tests, it is unclear why the 10% and 50% elutriate concentrations are tested in addition to the 100% concentration,

- when only the data from the 100% concentrations are used in the statistical analysis. There may be biological rationale for these additional concentrations of which I am unaware, such as action based on the lower dilution results should the 100% concentration be significantly toxic.

* There are a few places where statements about the test results do not match the data tables and graphs. These are listed below:

  1) In Section 3.3, page 9, it is stated that the ST4 sample was not toxic, while the ST3 sample may be toxic. The data tables and graphs indicate that ST3 is the nontoxic sample, while ST4 may be toxic.

  2) Section 5.5, page 22 states that there is "a significant difference between the control and weight of sample B2 30 percent treatment." The data analysis indicates that it is the sample B2 30 percent survival response which is significantly less than the control.

  3) Section 11.3, page 51 states that "survival in the group one samples was markedly higher than mysid survial in the group two samples." The data tables and charts indicate that the group two samples were high survival samples.

* A few pieces of data analysis output are missing from the document, particularly the sample B4 weight data analysis for the inland silverside stormwater test and the sample ST5 analysis for the polychaete sediment test.

While there exist alternative statistical techniques to those employed in this report that may have increased power to detect differences versus control groups, the analyses presented here are widely accepted and robust techniques for use with toxicity testings results. Please feel free to contact me with any questions regarding the specific comments provided.

cc: W. Thoeny, TAI

Appendix B

Relationship of Sample Stations with Other Site Features

APPENDIX B
Relationship of Sample Stations with Other Site Features

SEDIMENT STATIONS	MUSSEL STA-TIONS	STORM WATER STA-TIONS	BAY WATER STA-TIONS	OUTFALL AREAS	IR SITES	OU
S-1	M-1			B	IR-7, PA-18	IV
S-2	M-2			C	IR-6, IR-10	II
S-3	M-3	ST-1	B-1	D	IR-6, IR-10	II
S-4	M-4				IR-6	II
S-5	M-5	ST-2	B-2	G, H, I, J	IR-9	II
S-6	M-6				IR-8, IR-9	II
S-7	M-7				PA-16, IR-17	V
S-8	M-8			A	IR-11, IR-15, PA-16, IR-17	V
S-9	M-9				IR-2, IR-11, IR-15	I, V
S-10	M-10				IR-2, IR-3, IR-8, IR-11, IR-14, IR-15	I, II, V
S-11	M-11	ST-3	B-3		IR-2, IR-5, IR-12, IR-13	I, III, V
S-12	M-12				IR-2, IR-4, IR-5, IR-12	I, III, V
S-13	M-13				IR-1, IR-4	I, III
S-14	M-14				IR-1	I
S-15	M-15				Dry Dock #2/Dry Dock #3	
S-16	M-16	ST-4	B-4	E, F	S-203, S-209, S-210, S-215	
S-17	M-17				Dry Dock #4	

Appendix C

Summary Data Tables

APPENDIX C--Table 1
Summary of Stormwater Bioassay and Chemsitry Results

	ST-1	ST-2	ST-3	ST-4
CHEMISTRY	Pb>WQC Cu>WQC Zn>WAC Coliform	Pb>WQC Cu>WQC Zn>WQC Coliform	Pb>WQC Cu>WQC Mn>WQC Zn>WQC Coliform	Pb>WQC Cu>WQC Zn>WQC TCE Coliform
Acute Toxicity to Fathead Minnow	no	no	yes	yes
Chronic Toxicity to Fathead Minnow	no	no	no	no
Acute Toxicity to Ceriodaphnia*	inconclusive	yes*	inconclusive	yes*
Chronic Toxicity to Ceriodaphnia*	inconclusive	yes*	inconclusive	yes*
Chronic Toxicity to Selenastrum*	yes*	yes*	yes*	inconclusive

APPENDIX C--Table 2
Summary of Bay Water Bioassay and Chemistry Results

	B-1	B-2	B-3	B-4
Chemistry	none>WQC	none>WQC	none>WQC	none>WQC
Acute Toxicity to Inland Silverside	no	no	yes*	no
Chronic Toxicity to Inland Silverside	no	no	no	no
Chronic Toxicity to Purple Sea Urchin	no	no	no	no
Chronic Toxicity to Skeletonema	yes*	yes*	yes*	inconclusive

APPENDIX C--Table 3
Sediment Bioassay and Chemistry Results

	S-1	S-2	S-3	S-4
Exceeds ER-M at the surface	Ni, pyrene	Ni, pyrene, 4,4'-DDT	Ni	Pb, Cu, Ni, pyrene, Dieldrin, Endrin, 4,4'-DDD
Exceeds ER-L at the surface	Hg, Ni, Dieldrin, 4,4'-DDE, Endrin, 4,4'-DDD, 4,4'-DDT	Hg, Ni, Zn, Dieldrin, 4,4'-DDE, Endrin, 4,4'-DDT	Hg, Ni, 4,4'-DDE, Endrin	Hg, As, Pb, Cr, Cu, Ni, Zn, phenanthrene, fluoranthene, benzo(a)anthracene, chrysene
Exceeds mean of reference area at surface	As, Mn, DBT, TBT, delta-BHC, Dieldrin, 4,4'-DDD, methylene chloride	Ba, Cr, Fe, Mn, Ni, Zn, DBT, TBT, delta-BHC, 4,4'-DDE, 4,4'-DDD, 4,4'-DDT	As, Mn, DBT, TBT	As, Ba, Cr, Cu, Fe, Mn, Ni, Zn, MBT, DBT, TBT, Aldrin, Dieldrin, Endrin, 4,4'-DDD, alpha-Chlordane, Aroclor-1260
Exceeds ER-M at 2.5 feet	Ni, pyrene, Dieldrin	Ni, pyrene	Ni, pyrene	Hg, Cu, Ni, pyrene, Dieldrin, Endrin, 4,4'-DDE
Exceeds ER-L at 2.5 feet	Hg, Cr, Co, Ni, Zn, 4,4'-DDE, Endrin, 4,4'-DDD	Hg, Pb, Cr, Ni, Zn, Dieldrin, 4,4'-DDE, Endrin, 4,4'-DDD, 4,4'-DDT	Hg, Cr, Ni, Zn, 4,4'-DDE, Endrin, 4,4'-DDD	Hg, Pb, Cr, Cu, Ni, Zn, phenanthrene, fluoranthene, benzo(a) anthracene, chrysene
Exceeds mean of reference area at 2.5 feet	Ba, Cr, Co, Cu, Fe, Mn, Ni, V, DBT, TBT	Hg, Pb, Al, Ba, Cr, Cu, Fe, Mn, Ni, V, DBT, TBT, methyl ethyl ketone	Hg, Al, Cu, Fe, Mn, Ni, Zn, DBT, TBT, toluene, pyrene, 4,4'-DDE, Endrin, 4,4'-DDD	Hg, Pb, Ba, Cr, Cu, Fe, Mn, Ni, V, Zn, fluoranthene, pyrene, benzo(b)fluoranthene, benzo(a)pyrene, benzo(g,h,i) perylene, Endosulfan I, Dieldrin, Endrin, 4,4'-DDD, alpha-Chlordane, gamma-Chlordane
Acute toxicity to mysid* (elutriate)	yes*	yes*	yes*	inconclusive
Acute toxicity to sanddab* (elutriate)	inconclusive	inconclusive	inconclusive	yes*
Acute toxicity to bivalves (elutriate)	yes*	yes*	no	no

APPENDIX C--Table 3
Sediment Bioassay and Chemistry Results

Chronic toxicity to bivalves (elutriate)	yes	no	yes	no

APPENDIX C--Table 3
Sediment Bioassay and Chemistry Results

	S-5	S-6	S-7	S-8
Exceeds ER-M at the surface	Ni	Ni, pyrene, 4,4'-DDT	Ni, pyrene	Ni, pyrene, Indeno(1,2,3-cd)pyrene
Exceeds ER-L at the surface	Hg, Cr, Ni, Zn, 4,4'-DDE, Endrin	Hg, Ni, fluoranthene, chrysene, 4,4'-DDE, 4,4'-DDD	hg, Pb, Ni, phenanthrene, fluoranthene, benzo(a)anthracene, chrysene, 4,4'-DDE, Endrin, 4,4'-DDD	Hg, Ni, phenanthrene, fluoranthene, benzo(a)anthracene, chrysene, 4,4'-DDD
Exceeds mean of reference area at surface	Al, Ba, Cr, Fe, Mn, Ni, V, Zn, DBT, TBT, alpha-BHC, 4,4'-DDE, Endrin	Hg, As, Mn, DBT, TBT, alpha-BHC, beta-BHC, delta-BHC, 4,4'-DDT, Aroclor-1260	DBT, TBT	As, Mn, DBT, TBT, delta-BHC
Exceeds ER-M at 2.5 feet	Ni	Ni, Zn, pyrene	Pb, Ni, pyrene, 4,4'-DDT	Ni, pyrene, 4,4'-DDT
Exceeds ER-L at 2.5 feet	Hg, Cr, Ni, Zn, 4,4'-DDE, Endrin, 4,4'-DDD	Hg, Cr, Ni, Zn, Endrin	Hg, Pb, Ni, phenanthrene, 4,4'-DDE, Endrin, 4,4'-DDD	Hg, Ni, Zn, 4,4'-DDE, Endrin, 4,4'-DDD
Exceeds mean of reference area at 2.5 feet	Hg, Al, Ba, Cr, Co, Cu, Fe, Mn, Ni, V, Zn, DBT, TBT, acetone, toluene, delta-BHC, 4,4'-DDE, Endrin, 4,4'-DDD	Hg, Al, Ba, Cu, Fe, Mn, Ni, acetone, carbon disulfide, methyl ethyl ketone, pyrene, delta-BHC, methoxychlor	Pb, Cu, Mn, DBT, TBT, pyrene, 4,4'-DDE, 4,4'-DDD	Al, Ba, Be, Cr, Cu, Fe, Mn, Ni, Zn, DBT, TBT, toluene, pyrene, Dieldrin, acetone
Acute toxicity to mysid* (elutriate)	yes*	yes*	yes*	inconclusive
Acute toxicity to sanddab* (elutriate)	inconclusive	yes*	inconclusive	yes*
Acute toxicity to bivalves (elutriate)	yes*	no	no	no

APPENDIX C--Table 3
Sediment Bioassay and Chemistry Results

Chronic toxicity to bivalves (elutriate)	yes	yes	no	yes

APPENDIX C--Table 3
Sediment Bioassay and Chemistry Results

	S-9	S-10	S-11
Exceeds ER-M at the surface	none	Ni, pyrene, 4,4'-DDT	Ni, 4,4'-DDT
Exceeds ER-L at the surface	Ni, Dieldrin, 4,4'-DDE, Endrin, 4,4'-DDT	Ni, Dieldrin, 4,4'-DDE, Endrin, 4,4'-DDD	Hg, Cr, Ni, Zn, Dieldrin, 4,4'-DDE, Endrin, 4,4'-DDD
Exceeds mean of reference area at surface	Mn, alpha-BHC, gamma-BHC, Dieldrin, 4,4'-DDE, Endrin, 4,4'-DDT, methoxychlor, Aroclor-1260	Cu, Fe, Mn, Ni, V, Zn, DBT, TBT, alpha-BHC, Aldrin, Dieldrin, 4,4'-DDE, Endrin, 4,4'-DDD, 4,4'-DDT, methoxychlor, alpha-Chlordane, gamma-Chlordane, Aroclor-1260	As, Ba, Cr, Cu, Fe, Mn, Ni, Zn, DBT, TBT, alpha-BHC, gamma-BHC, Dieldrin, 4,4'-DDE, Endrin, 4,4'-DDD, 4,4'-DDT, alpha-Chlordane, gamma-Chlordane, Aroclor-1260
Exceeds ER-M at 2.5 feet	Pb, Ni, 4,4'-DDT	none	Ni
Exceeds ER-L at 2.5 feet	Hg, Pb, Cr, Ni, Zn, DBT, TBT, Endrin	Hg, 4,4'-DDE, Endrin	Hg, Cr, Ni
Exceeds mean of reference area at 2.5 feet	Hg, Pb, Al, Ba, Cr, Cu, Fe, Mn, Ni, V, Zn, DBT, TBT, Endrin, 4,4'-DDT, Aroclor-1260	Hg, Al, Ba, V, alpha-BHC, 4,4'-DDE, methoxychlor	Al, Cu, Fe, Mn, Ni, V, DBT, TBT, carbon disulfide
Acute toxicity to mysid* (elutriate)	yes*	inconclusive	inconclusive
Acute toxicity to sanddab* (elutriate)	yes*	yes*	inconclusive
Acute toxicity to bivalves (elutriate)	no	yes*	no
Chronic toxicity to bivalves (elutriate)	no	yes	no

APPENDIX C--Table 3
Sediment Bioassay and Chemistry Results

	S-12	S-13	S-14
Exceeds ER-M at the surface	Pb, Ni, pyrene, Dieldrin, 4,4'-DDD, 4,4'-DDT	Ni, Dieldrin, 4,4'-DDE, 4,4'-DDD, 4,4'-DDT	Pb, Ni, pyrene, Dieldrin, 4,4'-DDT
Exceeds ER-L at the surface	Hg, Pb, Cr, Cu, Ni, Zn, 4,4'-DDE	Hg, Pb, Cr, Cu, Ni, Zn, Endrin	Hg, Pb, Cr, Cu, Ni, Zn, 4,4'-DDE, Endrin, 4,4'-DDD
Exceeds mean of reference area at surface	Hg, Ba, Cr, Cu, Fe, Mn, Ni, V, Zn, DBT, TBT, Dieldrin, 4,4'-DDE, 4,4'-DDD, 4,4'-DDT, alpha-Chlordane	Hg, As, Ba, Cr, Cu, Mn, Ni, Zn, DBT, TBT, Dieldrin, 4,4'-DDE, Endrin, 4,4'-DDD, 4,4'-DDT, alpha-Chlordane gamma-Chlordane, Aroclor-1260	Al, Ba, Cr, Cu, Fe, Mn, Ni, V, Zn, DBT, TBT, Dieldrin, 4,4'-DDE, Endrin, 4,4'-DDD, 4,4'-DDT, alpha-Chlordane, gamma-Chlordane, Aroclor-1260
Exceeds ER-M at 2.5 feet	Pb, Cr, Ni, pyrene	Pb, Cr, Ni, Zn, pyrene, Dieldrin, 4,4'-DDE, 4,4'-DDD	Cr, Ni, pyrene, Dieldrin, 4,4'-DDT
Exceeds ER-L at 2.5 feet	Hg, Pb, Cr, Cu, Ni, Zn, Dieldrin, 4,4'-DDE, Endrin, 4,4'-DDD, 4,4'-DDT	Hg, Pb, Cr, Cu, Ni, Zn, fluoranthene, chrysene, Endrin	Hg, Pb, Cr, Cu, Ni, Zn, 4,4-DDE, Endrin, 4,4'-DDT
Exceeds mean of reference area at 2.5 feet	Hg, Pb, Al, Ba, Be, Cr, Cu, Fe, Mn, Ni, V, TBT, Dieldrin, 4,4'-DDE, Aroclor-1260, carbon disulfide, methyl ethyl ketone	Hg, Al, Ba, Cr, Cu, Fe, Mn, Ni, V, Zn, DBT, TBT, fluoranthene, chrysene, Dieldrin, 4,4'-DDE, Endrin, 4,4'-DDD, alpha-Chlordane, gamma-Chlordane, acetone, carbon disulfide, methyl ethyl ketone	Hg, Pb, Al, Ba, Cr, Co, Cu, Fe, Mn, Ni, V, Zn, DBT, pyrene, Dieldrin, 4,4'-DDE, Endrin, 4,4'-DDT, Aroclor-1260, toluene
Acute toxicity to mysid* (elutriate)	yes*	inconclusive	yes*
Acute toxicity to sanddab* (elutriate)	yes*	inconclusive	inconclusive
Acute toxicity to bivalves (elutriate)	no	yes*	no

APPENDIX C--Table 3
Sediment Bioassay and Chemistry Results

Chronic toxicity to bivalves (elutriate)	yes	yes	yes

APPENDIX C--Table 3
Sediment Bioassay and Chemistry Results

	S-15	S-16	S-17
Exceeds ER-M at the surface	Pb, Ni, pyrene	Pb, Ni, pyrene	Ni, 4,4'-DDT
Exceeds ER-L at the surface	Hg, Pb, Ni, 4,4'-DDE, Endrin,	Hg, Pb, Cr, Ni, Zn, 4,4'-DDE, Endrin	Hg, Ni, Dieldrin, 4,4'-DDE, Endrin,
Exceeds mean of reference area at surface	As, Ba, Cr, Co, Fe, Mn, Ni, DBT, TBT, Endrin	As, Cr, Cu, Fe, Mn, Ni, Zn, DBT, TBT	As, Al, Ba, Cr, Fe, Mn, Ni, TBT, 4,4'-DDT
Exceeds ER-M at 2.5 feet	Hg, Ni, pyrene, 4,4'-DDE, 4,4'-DDT,	Pb, Ni, pyrene, 4,4'-DDT	Pb, Ni
Exceeds ER-L at 2.5 feet	Hg, Pb, Cr, Ni, Zn, 4,4-'DDE, 4,4'-DDD	Hg, Pb, Cr, Ni, Zn, Dieldrin, 4,4'-DDE, Endrin	Hg, Pb, Cr, Ni, Zn, 4,4'-DDE, Endrin, 4,4'-DDD
Exceeds mean of reference area at 2.5 feet	Hg, Pb, Al, Ba, Cr, Co, Cu, Fe, Mn, Ni, V, DBT, TBT, pyrene, Dieldrin, 4,4'-DDE, Endrin, 4,4'-DDT, Aroclor-1260, toluene	Hg, Pb, Al, Ba, Cr, Co, Cu, Fe, Mn, Ni, V, DBT, TBT, pyrene, Dieldrin, 4,4'-DDE, Endrin, 4,4'-DDT, gamma-Chlordane, acetone, toluene	Pb, Al, Ba, Cu, Fe, Mn, Ni, V, DBT, TBT, 4,4'-DDE, 4,4'-DDD, methyl ethyl ketone
Acute toxicity to mysid* (elutriate)	yes*	yes*	yes*
Acute toxicity to sanddab* (elutriate)	inconclusive	yes*	inconclusive
Acute toxicity to bivalves (elutriate)	yes*	yes	no
Chronic toxicity to bivalves (elutriate)	yes	yes	no

* Those results which were either invalid or inconclusive have been reviewed more broadly and the results therefore marked with an asterik .

APPENDIX C--Table 4
Summary of Mussel Tissue Chemistry Results

	M-1	M-2	M-3	M-4	M-5	M-6
Exceeds mean of reference area	Al, Sb, As, Ba, Cd, Cr, Co, Cu, Fe, Pb, Mn, Mo, Ni, Se, Ag, Sn, V, Zn, DBT, TBT, phenanthrene, fluoranthene, pyrene	Al, Sb, As, Ba, Cd, Cr, Co, Cu, Fe, Pb, Mn, Hg, Mo, Ni, Se, Ag, Sn, V, Zn, MBT, DBT, TBT, diethyl phthalate	Al, Sb, As, Ba, Cd, Cr, Co, Cu, Fe, Pb, Mn, Hg, Mo, Ni, Se, Sn, V, Zn, MBT, DBT, TBT, diethyl phthalate, phenanthrene, fluoranthene, pyrene	Al, Sb, As, Ba, Cd, Cr, Co, Cu, Fe, Pb, Mn, Hg, Mo, Ni, Se, Ag, Sn, V, Zn, MBT, DBT, TBT	Al, As, Cd, Pb, Mo, Se, Sn, DBT	Al, As, Ba, Cd, Cr, Co, Cu, Fe, Pb, Pb, Mn, Mo, Ni, Se, Ag, Sn, V, Zn, MBT, DBT, TBT

	M-7	M-8	M-9	M-10	M-11	M-12
Exceeds mean of reference area	Al, As, Ba, Cd, Cr, Co, Cu, Fe, Pb, Mn, Mo, Ni, Se, Ag, Sn, V, DBT, TBT	Al, Sb, As, Ba, Cd, Cr, Co, Cu, Fe, Pb, Mn, Hg, Mo, Ni, Se, Ag, Sn, V, Zn, TBT	Al, As, Ba, Cd, Cr, Cu, Fe, Mn, Ni, Se, V, Zn, DBT, TBT, 4,4'-DDE, Aroclor-1254	Al, Sb, As, Ba, Cr, Cu, Fe, Pb, Mn, Mo, Ni, Se, Ag, Sn, V, MBT, TBT, 4,4'-DDE, Aroclor-1260	Al, As, Ba, Cd, Cr, Co, Cu, Fe, Pb, Mn, Hg, Mo, Ni, Se, Sn, V, Zn, DBT, TBT, 4,4'-DDE, Aroclor-1254	Al, As, Ba, Cd, Cr, Co, Cu, Fe, Mn, Hg, Ni, Ag, Sn Zn, TBT, 4,4'-DDE, Aroclor-1260

	M-13	M-14	M-15	M-16	M-17
Exceeds mean of reference area	Al, As, Cd, Cr, Cu, Pb, Mn, Ag, Sn, V, TBT, 4,4'-DDE, Aroclor-1260	Al, Sb, As, Ba, Cd, Cr, Co, Cu, Fe, Pb, Mn, Hg, Mo, Ni, Se, Ag, Zn, TBT, 4,4'-DDE, Aroclor-1254	Al, Sb, As, Ba, Cd, Cr, Co, Fe, Mn, Mo, Ni, Se, Ag, Sn, V, DBT, TBT, phenanthrene, fluoranthene, pyrene, 4,4'-DDE	Al, Sb, As, Ba, Cd, Cr, Cu, Fe, Mn, Mo, Ni, Se, Sn, V, DBT, TBT	Sb, As, Ba, Cr, Cu, Fe,Mn, Mo, Ni, Se, V, DBT, TBT

ATTACHMENT 2

NOAA'S ESAP REVIEW COMMENTS

U.S. DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
NATIONAL OCEAN SERVICE
OFFICE OF OCEAN RESOURCES CONSERVATION AND ASSESSM
HAZARDOUS MATERIALS RESPONSE AND ASSESSMENT DIVIS
COASTAL RESOURCES COORDINATION BRANCH
c/o U.S. Environmental Protection Agency (H-1-2)
75 Hawthorne Street
San Francisco, CA 94105-3901

August 20, 1993

Ms. Alydda Mangelsdorf
USEPA
75 Hawthorne Street (H-9-2)
San Francisco, CA 94105

Dear Alydda:

The U.S. Department of Commerce/National Oceanic and Atmospheric Administration (NOAA) appreciates the opportunity to review the Supplemental ESAP Data Submittal and Data Validation Summary Analysis for Naval Station Treasure Island, Hunters Point Annex, San Francisco, CA. These documents were prepared by PRC Environmental Management, Inc., San Francisco, and Harding Lawson Associates, Novato, CA and submitted for review on April 15, 1993.

These documents consist of data collected according to the Environmental Sampling and Analysis Plan (ESAP) of July 31, 1991. As of this submittal, no interpretation of the results, with respect to the potential threat posed to aquatic organisms, was included in the data. In addition to the sediment, mussel tissue, and bioassay data from the ESAP, data from sediment sampling from IR01, 02, 03, and 07, collected between 1/1/88 and 12/31/92, was included in the submittal.

Summary--Bioassay Results

Bioassays were conducted on stormwater discharge using a 7-day fathead minnow (*Pimephales promelas*) assay, 7-day cladoceran (*Ceriodaphnia dubia*), and a 96-hour green algae (*Selenastrum capricornutum*). Out of the three bioassays using fresh water species, data were considered valid only for the fathead minnow bioassay. In the fathead minnow bioassay, two stormwater stations (ST1 and ST2) did not differ significantly from the control. There was significant reduction in survival at 100% sample concentration for stations ST3 and ST4. Conclusions could not be drawn from the other two bioassays due to confounding factors affecting the assays that were not related to sample dose.

Bioassays were also conducted on stormwater discharge using a 7-day inland silverside (*Menidia beryllina*), a 1-hour purple sea urchin (*Strongylocentrus purpuratus*), and a 96-hour diatom (*Skeletonema costatum*). Results from the silverside bioassay indicated that the data were valid and the stormwater did not show a statistically significant affect on survival or growth. The urchin assay showed toxicity, and the diatom study was inconclusive due to testing problems.

In the sediment elutriate bioassays, data quality conformance standards were not met in the 96-hour mysid shrimp (*Holmesimysis costata*), and the 96-hour sanddab (*Citharicthys stigmaeus*) assays. The 48-hour Pacific oyster (*Crassostrea gigas*) data were

valid and resulted in mortality and developmental abnormalities. Although the mortality was not considered statistically different from the controls, the developmental abnormalities were higher than those occuring in the controls.

For the bulk sediment bioassays, the 10-day mysid shrimp (*Holmesimysis costata*) results were inconclusive with mortality occurring in group one and no significant mortality occurring in group two. The other two bioassays, the10-day amphipod (*Eohaustorius estuarius*), and the10-day polychaete (*Nephtys caecoides*) resulted in mortality but the results are considered inconclusive due to other laboratory factors.

Summary--Sediment and Mussel Tissue Results

Maximum contaminant levels reported for sediment at the IR sites show metals elevated at all sites, with the highest levels found at IR01 and IR02. These levels in sediment at IR01 and 02 exceed the NOAA Effects Range-Low (ER-L) and the EPA Region 10/Puget Sound Apparent Effects Threshold (AET) levels. Maximum PAH levels were elevated above the ER-L for IR01 and above the AET levels for selected PAHs. For the ESAP data, maximum levels of metals, PAHs, PCBs, and pesticides tended to be lower than the levels reported for IR01 and 02, possibly due to compositing of the sediment. However, the ESAP levels still exceed the ER-L and the AET for several metals, PAHs, PCBs, and pesticides.

Results from the mussel tissue data indicate that tributyltin, metals, PAHs, DDE, and PCBs have bioaccumulated in mussels at Hunters Point.

Comments

The results from eight of the bioassays were inconclusive due to poor data quality and possible methodology problems. The tests that were considered valid did show an adverse effect on biota from contaminants in the sediments. Although the valid tests indicate that the site poses a threat to aquatic organisms, the data are not adequate to perform the ecological risk assessment that was originally described in the September 9, 1992 document.

Based on the September 9, Draft Final Ecological Risk Assessment Work Plan, a Phase1B sampling effort will be conducted to determine the potential impact of site related contaminants to fish and invertebrates. According to the work plan, in addition to the activities described, the results of the Phase 1A effort will be used to design and implement additional field sampling and analyses plans for subsequent phases of work. Levels of contaminants found in the sediments indicate a need to do a further assessment of the lateral and vertical extent of contaminated sediments. Information on the lateral and vertical extent of contaminated sediments will be useful in determining where to focus future field sampling and analysis, as well as provide information for remedial option decision making.

Considering the levels of contaminants in the sediment, the bioaccumulation of metals in mussel tissue, and the lack of conclusive data for most of the bioassays, NOAA recommends that the Navy conduct additional studies to further assess survival, growth

and developmental impacts on biota that may be resulting from past releases of contaminants at Hunters Point Annex. The documentation submitted with the results from future studies should include a description of the physical parameters measured during the bioassays (e.g. pH, oxygen, temperature), as well as characteristics such as total organic carbon (TOC), total sulfide, and sediment grain size for later interpretation of the results. Chemical analysis should be performed on the sediment samples that are used for the bioassays so that the toxicity determined in the bioassays can be correlated to the concentration of the contaminants.

The studies planned for phase 1B will provide additional information on impacts to biota and should be conducted. However, the results of phase 1B are intended to supplement the bioassay data and should not be used in place of bioassay data for evaluating the ecological risk to aquatic biota at Hunters Point.

If you have any questions about these comments or require further explanation or elaboration, I may be reached at (415) 744-3126.

Sincerely,

Denise M. Klimas

Denise M. Klimas
Coastal Resources Coordinator

cc: Ms. Roberta Blank, EPA
Dr. Barbara Smith, RWQCB
Mr. Cyrus Shabahari, DTSC