Naval Training Center, San Diego
Restoration Advisory Board Comments on Draft Work Plan
Sediment Characterization of the NTC Boat Channel
CTO-092 Response to RAB Comments

David Valentine
Specific Comment 1:
Page 1-1, Introduction, last two paragraphs. The "tone" for the field investigation is set here as *"to identify the presence and general spatial distribution of chemical constituents" "to better characterize the vertical and horizontal distribution and dynamics of sediment contaminants within the Boat Channel."*

Response to Specific Comment 1:
Comment noted.

Specific Comment 2:
Page 2-1, 4th paragraph. *"The Boat Channel is the last remnant of the original river bed of the San Diego River."* This is probably true. Note, however, that the channel was dredged to its present configuration in 1946. Whether this channel, as it now exists, follows the last river course is problematical. Regardless, dredging did occur and this would have removed the upper layers of riverine sediment. Any additional inputs would, then, have been laid down after the cessation of hostilities in WWII and after the introduction of most organic toxicants of concern. A review of historic aerial photographs should determine how much of the channel follows its historic bed.

Response to Specific Comment 2:
Comment noted. Aerial photos of the Boat Channel will be reviewed for the final report.

Specific Comment 3:
Page 2-1, 5th paragraph, 6th line. *"Significant flushing action from San Diego Bay does not occur."* What is viewed as "significant"? What type of interchange is needed?

Response to Specific Comment 3:
The term "significant flushing" was used as it was determined by Luke-Dudek (1983) page V-17 of their study. They reported that there are "considerable turbulence and eddy conditions" at the north end of the channel and at the western tip of Harbor Island. The report stated that waters in the northern end tended to rotate in a circular pattern indicating poor flushing capacity. Water in the main channel moved back and forth with the tides with little intermixing with Bay water. The turbulence near the Harbor Island inlet suggested that Bay water tends to flow into the Harbor Island Boat Basin instead of the channel. The term "significant" was interpreted to mean that there was no clear-cut or discernible exchange of waters into and out of the Boat Channel that could be detected during the course of the drogue study.

Specific Comment 4:
Page 2-3, Figure 2-2. Define "FFTA"

Response to Specific Comment 4:
Between 1944 and the 1960's fire fighting training was conducted at Buildings 401, 4996, and 555, with are located at the Former Firefighter Training Area (FFTA). During training exercises, materials such as waste fuels and shop wastes were burned on the pavement. Unknown volumes of water used to quench the fires were drained to San Diego Bay.

Specific Comment 5:
Page 2-7, Figure 2-4. The tidal cycle should be superimposed on this figure. This is supposed to be a "flood" drogue study but there is no indication of when the slack tides occurred nor the tidal excursion for the time period during which the study occurred. Even a casual inspection of this figure will demonstrate

that the drogues in the channel went to the north but the southern most drogue went south. This is counterintuitive. Also, the drogues in the mouth of the channel exhibit a complex pattern which cannot easily be explained by supposing that they were dropped on a flood tide.

Response to Specific Comment 5:
The tidal cycle has been added to Figure 2-4; this figure shows drogue movement from a low slack tide to a high tide. The Luke-Dudek study (1983) did not specify slack points or lag times for various areas in San Diego Bay. The report attributes the drogue movements to turbulence areas between water bodies. Also, it is unknown what influence boat traffic may have had on these measuring devices.

Specific Comment 6:
Page 2-8, Figure 2-5. The same comments are applicable to this figure as on the previous one. According to this figure, drogues released at or near the end of Harbor Island always go into the Bay. This is clearly impossible. Also, the two drogues released in the entrance of the Boat Channel moved toward the east while the drogue supposedly in the east of the Harbor Island Channel moved west, and then presumable [sic] toward the Bay. Interesting observation, if true.

Response to Specific Comment 6:
It must be kept in mind that this was a one-time event. The flood tide drogue nearest the San Diego Bay may have re-circulated into the channel. Regarding the two channel drogues, there often are differential counter-currents along the edges of waterways.

Specific Comment 7:
Page 2-11, last paragraph, last line. "*Sediment chemistry values falling within the lower tenth percentile concentration range (effects range-low [ERL]) and the fiftieth percentile concentration (effects range-median [ERM]) were associated with observed or predicted effects.*" This is a rather positive statement which is counter-intuitive. If biological effects were manifested in the lower 10th percentile range then one need not predict effects for higher ranges - either they would occur or the testing protocol would have been faulty.

It would seem that the manner in which probable effect was defined is purely statistical and has nothing to do with observed effects. That is, one obtains a data set and calculates the appropriate percentiles. Straightforward. However, assuming that these data have any biological meaning is dangerous.

If, in fact, what was described is a purely exercise (this is certainly intimated by the discussion), then it is absolutely impossible to make a statement such as, "*All metals concentrations were below the ERM.*" This is clearly impossible since one cannot take a data set, calculate a mean, and then have all of the data below the mean! Perhaps a rethinking of this discussion is warranted.

Response to Specific Comment 7:
Because there are no specific sediment quality guidelines for marine sediments in California, let alone San Diego Bay, the widely-used default values for sediment are the ERMs and ERLs as described in Section 2 of the Draft Work Plan. These values were developed by the National Oceanic and Atmospheric Administration (NOAA) have been recommended by both the state and federal agencies as guidelines when they have needed to provide a numerical screening value. To clarify, the database from which the ERMs and ERLs were derived include datasets from modeling (e.g., equilibrium partitioning), laboratory (e.g., bioassay and chemistry analyses), and field studies (e.g., apparent effects threshold, sediment triad approach). Thus the dataset contains both biological and chemical data, somewhat similar to the manner in which water quality guidelines are developed.

Concerning the statement "*All metals concentrations were below the ERM.*" The data that are referenced in this statement are single data points; no means were calculated and compared against the ERLs and ERMs. It must be understood that the BPTCP data were not used in developing the ERLs and ERMs.

For further discussion and details concerning the ERL and ERM values, we recommend the following references:

Long, E.R., D.D. MacDonald, S.L. Smith, and F.D. Calder. 1995. Incidence of Adverse Biological Effects Within Ranges of Chemical Concentrations in Marine and Estuarine Sediments. Environmental Management Vol. 19, No. 1, pp. 81-97.

Long, E.R. and L.G. Morgan. 1990. The Potential for Biological Effects of Sediment-Sorbed Contaminants tested in the National Status and Trends Program. NOAA Tech. Memo. NOS OMA 52. U.S. National Oceanic and Atmospheric Administration, Seattle, Washington, 175 pp.

Specific Comment 8:
Page 1-12, 2.3.2, Toxicity Results [sic; this should read page 2-12]. The data set being described seem substandard since one cannot make inferences about biological effects unless chemical and biological data are obtained on the same sediments.

Response to Specific Comment 8:
We concur with the comment. For the BPTCP, toxicity and chemistry data were not collected concurrently on all samples. Because the BPTCP data set appears to be somewhat incomplete, this study was needed to provide more information for the Boat Channel sediments.

Specific Comment 9:
Page 3-2, section 3.3.1, Sampling Strategy. The definition of "strata" is questioned. One establishes strata ***IF*** there are recognizable differences between or among strata (chemically, biologically, topographically, hydrologically, etc). No such differences were established. Assuming "significant" hydrological differences based on the data previously presented is dangerous. Even a simple inspection of the drogue data suggests that there is considerable interchange of water. Where does one establish the "strata" with these data?

Response to Specific Comment 9:
Strata in this case have been established based on review of existing data, discussions with regulators, discussions with statisticians, and best professional judgment of the technical specialists. Partitioning the channel into three separate areas or zones based on hydrology (i.e., apparent circulation) was considered to be a reasonable approach for locating sampling stations.

Specific Comment 10:
Page 3-7, first paragraph. Placing the three samples in each strata [sic] randomly is also questioned. Locating the sampling locations using this method will allow you to calculate the likely *mean* concentration of toxicants in each "strata", but will minimize the amount of information gained regarding the horizontal distribution. The introductory section states that the aims of the study were "*to identify the presence and general spatial distribution of chemical constituents*" "*to better characterize the vertical and horizontal distribution and dynamics of sediment contaminants within the Boat Channel.*"

The first aim cannot be optimally achieved using a random sampling pattern. Little can be gained on the "dynamics" using this approach. A more reasonable approach would be to place the samples evenly from the upper to the lower boat channel. this gives one the ability to compare and contrast differences between or among one or more data sets and also gain a better perspective into the horizontal distribution of contaminants.

Response to Specific Comment 10:
Spacing the sampling locations evenly along the length of the Boat Channel was discussed with two different statisticians within BNI and SWDIV. Both felt that the power of the statistics would be lost using a judgmental approach as suggested in this comment. Development of future sampling events, should they be deemed necessary, will have greater usage if the original sampling event has a degree of randomness

associated with it. According to U.S. EPA, randomization is used to make probability or confidence statements about data. Results from judgmental sampling cannot be generalized to the entire sample area and no probability statements can be made.

A drawback of judgmental sampling is that it is difficult to measure the accuracy of the estimated parameters because information on the variability of the data cannot be assured using statistical techniques. As a result, judgmental sampling may yield accurate results, but the degree of accuracy cannot be quantified with confidence (USACE, 1994: Technical Project Planning Guidance for HTRW Data Quality design, Draft Final, U.S. Army Corps of Engineers, April 1994).

Specific Comment 11:
Page 3-7 Sample Collection. Arbitrarily dividing the cores into six subsections is also questioned. Toxicant distribution in sediments is more closely tied to stratification than arbitrary subdivision. That is, the cores may show varving (layering). It would be far more applicable to test each of the layered sections for chemical and toxicological effects. Doing otherwise may well blur distinctions between adjacent core segments. For instance, your basic hypothesis is that the upper cores will be more contaminated than the lower cores. Somewhere in the core length there may well be a varve (this is typical of other sediment studies conducted in San Diego Bay and elsewhere). It is conceivable that contaminated (upper) sediments might be mixed with clean (lower) sediments which would blur the data derived fro these studies. It is likely better to test sediments based on sediment characteristics than on a fixed vertical distribution. It would also likely be far less expensive since is it [sic] viewed as unlikely there will be six strata in these cores.

Response to Specific Comment 11:
Comment noted. The sampling approach is based on agency recommendations on consistency between sample cores, specifically, the agencies prefer to see sediment layers that are directly comparable in "distance" rather than geologic strata. The set number of subsections is also somewhat based on remediation technology. If removal were necessary, it currently can only be accomplished with a two foot accuracy. Should remediation be found to be necessary, removing individual strata (varves) at various depths along the channel may not be appropriate for remediation or technically feasible.

Specific Comment 12:
Page 3-11, 4th paragraph. *"For all of the toxicity tests, sediments from which the test organism originated will be tested along with sediments from the Boat Channel"*. It appears that all of the sediments will be from the boat channel. How does this "negative" control differ from the other sediments to be tested?

"Positive control" sediments for dredging bioassay work have historically be collected off of Mission Bay. where are these control sediment [sic] to be taken from? Some past bioassay work (for the Commercial Basin, for instance), were taken from the main boat channel. These main boat channel sediments were often found more toxic (demonstrated more negative biological effects) than the sediments being tested. The sediment control area is, thus, very important.

Response to Specific Comment 12:
Negative control sediments are sediment from which the test organisms originated; they are not Boat Channel sediments. Depending on the organism and supplier, these negative control sediments may be from Tomales Bay, Newport Bay, or Mission Bay.

Positive control sediments are sediments spiked with a reference toxicant, such as copper or chromium compounds. Mission Bay sediments, on the other hand, can be reference sediments against which the dredging area sediment are compared to determine significant adverse effects. They may not necessarily be positive control sediments, depending on the focus of the study. We would be interested in obtaining the bioassay results from Commercial Basin, where the main boat channel sediments were used as "control" sediments.

Specific Comment 13:
4-1, Section 4.3 Sediment Sample Collection, last sentence. This makes reference to analyzing strata within a core. Consider introducing the concept discussed earlier.

It may be that the cores are being sectioned in the field to facilitate transportation to the laboratory. That is, it is easier to transport six one foot cores than one six foot core, or two three foot cores, or three two foot cores. Suffice it to say that the less the cores are handled in the field the less chance there is for the introduction of potential contaminants and the less chance that the cores will be inadvertently mislabeled.

Response to Specific Comment 13:
Comment noted. It is recognized that increased handling of cores in the field may increase the possibility of contamination. However, it is believed that it is more expedient to subsection in the field because there is less chance for inaccurate subsectioning in the laboratory and prevention of field contamination can be achieved.

Specific Comment 14:
Page 6-1, Data Analysis. The section is reasonably clear though, in my opinion, potentially flawed. an ANOVA does require randomly obtained samples to be truly valid. However, this violates the need to first determine the spatial distribution of contaminants within the boat channel. Remember, the primarily thrust of the investigation is to determine the vertical and horizontal distribution of contaminants and not necessarily if the strata are somehow "equivalent" to one another.

Also, one is only going to have three samples from each strata. Calculating a mean within a strata results in the loss of one degree of freedom. If another transformation is required to reach data normality than a second degree of freedom will be lost. This leave [sic] only one degree of freedom which results in an ANOVA which approaches powerlessness.

Response to Specific Comment 14:
Comment noted. It is recognized that the small sample size (n) per stratum can be considered problematic for statistical analyses. The costs of conducting this type of work - sampling, chemistry, and toxicological analyses - are very high and compared to groundwater and soil contaminant investigations. If, for example, chemistry or toxicity tests were conducted to the exclusion of the other, a greater coverage would be achieved, but data usage would be compromised. It is typical with sediment studies to sample over relatively large areas in order to characterize a site.

Specific Comment 15:
6th paragraph, 6th line. An ANOVA does require a normal distribution but is [sic] does not require a "constant" variance. What it does require is that the variances not be significantly different. In other words, that the data points being tested could have been drawn from the same population, not that they are constant. Note also that as the degrees of freedom are decreased that the power of Bartlett's test decreases significantly. With one degree of freedom the test borders on uselessness as does the other frequently used variance test, Snedecor's F test.

Response to Specific Comment 15:
Comment noted. The statistical tests described in Section 6.1 for the toxicity tests are routinely used to evaluate bioassay results from sediment evaluations.

Specific Comment 16:
Draft Field Sampling Plan. Page A2-1. Field measurements. Routinely measuring pH in seawater is not likely to tell you anything. It is a cheap, simple, tests but seawater is so highly buffered that massive quantities of acid or base would have to be injected to have any discernible impact on pH. Also, conductivity by itself is of questionable utility. The measurement of choice is salinity which, of course, can be obtained using conductivity and temperature readings. Thus, state that salinity will be calculated.

Response to Specific Comment 16:
The field measurements will be collected only for descriptive purposes and will not be used in any statistical analyses. Salinity readings will be presented in the summary report.

^^

C.B. Bishop
General Comment 1:
The following comments fall mostly in the "educational" category in the hope that the answers will enable the RAB to better understand the elements of the work plan.

Response to General Comment 1:
Comment noted.

Specific Comment 1:
2.1 Previous measurements of water exchange in the channel are suspect if they indicate little flushing action between the Bay and the channel. Note that a six foot tidal range equates to about bout 12 million cubic feet of water in the channel (6x400x5000). This usually happens several times in a month, e.g., ten times last October. This probably represents 30-50% water exchange, depending on channel depth, which is not stated. The report states that ground water flow direction is sometimes reversed by tidal action. This would indicate that the channel's fluid contents should have some significant effects from tidal action.

Response to Specific Comment 1:
The data presented in this section was obtained from a previous report by Luke-Dudek (1983). Re-interpretation of their assumptions and conclusions was not conducted. We do concur with the comment that waters within the channel can be significantly affected by tides.

Specific Comment 2:
Figs. 2-4 & 2-5 These figures are confusing. They don't tell what the tidal times are or the ranges. They have water moving in opposite directions, with flood and ebb tides in the same time periods. The don't give much confidence in the measurements or analysis.

Response to Specific Comment 2:
In small channels and embayments, there can be water movement opposite to that of the prevailing tide. There can be a lag time of some areas where the water may move at a different rate than the majority of the flow. Also, friction and drag around physical structures can cause local eddy patterns.

Specific Comment 3:
Tables 2-3,4,5,6 Better have someone prepared to explain these tables and ensuing toxicity discussions. How was it ensured that no other agents could have affect the specimens other than the test samples?? How do the numbers presented relate to regulatory standards?

Response to Specific Comment 3:
The tables present data prepared from the Bay Protection and Toxic Cleanup Program, run through a joint effort between the State Water Resources Control Board, Regional Water Quality Control Boards, and California Department of Fish and Game. To our understanding, the bioassays were conducted by CDFG personnel experienced with these types of tests and the tests were subjected to standard laboratory quality control measures. The manner in which the tests are performed ensure that the effects are a result of reactions to the test media (sediment or porewater). A complete report by the agencies is not yet available, so that a complete picture is not possible at this time. The numbers (i.e., results) do not relate to any regulatory standards, as there are none for characterization work.

Specific Comment 4:
Section 3.1 Will these tests account for all probable pollutants, or just those we think may have been deposited in the landfill or storm drains? for example, the residue from unburnt fuel and exhaust gases from departing aircraft?

Response to Specific Comment 4:
The analytical tests described in the Draft Work Plan cover the entire suite of chemicals that are usually tested, i.e., no chemicals have been screened based on results of previous studies of the storm drains or landfill. If any fuel components such as unburnt fuel or exhaust gases, are not water soluble and sorb to sediment, they would most likely be detected, if present in high enough concentrations.

Specific Comment 5:
Section 3.4.1 Please have someone explain the differences in the QC/QA standards, with respect to risk vs cost, if possible.

Response to Specific Comment 5:

An adequate response to this comment will require more details from the commentor. The various levels of QA/QC analysis and validation refer to the completeness of the data package. It is unclear what the term "risk" refers to in the commentors statement. Risk can be defined in terms of the risk of obtaining inadequate data; the risks described in a risk assessment; risks in terms of legal obligations, to name a few. Further clarification of this comment is needed in order to fully respond.

^^

Z Kripke, MD, MPH
General Comment 1:
But there will be only 10 very localized sites studied at one point in time and without any plan to map plumes or horizontal extent of any "hits," which does not fulfill this mandate to study the distribution, much less the dynamics of contamination.

Response to General Comment 1:
In contrast to studies of contaminated groundwater or soil, studies in the marine environment assume movement of potentially contaminated media due to tidal conditions and seasonal events, however the movement may not necessarily be in any one particular direction. The marine environment is in many respects far more dynamic than land-based sites. Therefore "plumes" as commonly mapped for groundwater are not mapped; usually relative gradients of sediment concentrations can be generally inferred without fine gradations shown. This study will not answer all of the questions regarding contamination and distribution because final use is not known, but it will act as a guide for further study as deemed necessary.

General Comment 2:

No matter how high the sediment chemistry numbers (P. 6-3), there will be no further action if sediment toxicity and bioaccumulation tests are negative. Yet in the past testing of the Bay Protection and Toxic Cleanup Program of 1994, there was essentially complete decoupling of the chemistry and toxicity for Rhepoxynius abronius [sic]. (Results show this lack of correlation repeatedly - for example P. 2-26 "Comparison of chemical and toxicological results from these two samples showed no correlation patterns").

Response to General Comment 2:

It should be understood that there will be a complex suite of toxicity tests conducted with many different species of marine organisms that have varying sensitivity. We will also be measuring multiple endpoints, e.g., survival, growth, development. The no further action decision will be based on the premise that all of the toxicity tests conducted on channel sediments are clearly not significant as compared to San Diego Bay sediments. Even if this result is observed, which is rare, there will still be extensive discussion as to the true meaning of these results.

Specific Comment 1:

Since one goal is to characterize the horizontal distribution of contaminants, how will the footprint of any positive samples be mapped?

Response to Specific Comment 1:

One of the results of this study will be to provide an overall snapshot of contaminant distribution throughout different parts of the channel. "Footprints" at this scale of sampling cannot be identified; general inferences about gradients may be identified. Graphical representations of the sediment contaminants will be presented in the summary report.

Specific Comment 2:

Since another goal is to characterize the dynamics of the spread of sediment contaminants, how will this be mapped, plumes outlined, etc.?

Response to Specific Comment 2:

See Response to Specific Comment 1.

Specific Comment 3:

Since past experience is that toxicity and chemistry are not correlated, positive chemistry levels must be investigated further for footprint, movement and origin rather than listed as 'no further action' (p. 6-3).

Response to Specific Comment 3:

This a broad assumptions by the commentor that toxicity and chemistry are not correlated. The associations as observed in the Bay Protection and Toxic Cleanup Program (BPTCP) data may not have been clear-cut, but it is not advisable to assume that all bioassay results are thus related. As stated previously, this is a characterization study and the results will be used to assist in determining further actions.

Specific Comment 4:

If soils and waters prove highly toxic to organisms but the chemicals currently studied do not correlate with this toxicity, a search must be conducted for what is poisoning the organisms. The chemicals being tested for should be expanded to include any reasonable hypotheses or to test synergisms among the chemicals present that might account for high mortality in some sediments and waters where each contaminant is at the ERL.

Response to Specific Comment 4:

It is recognized that there may be synergism among some sediment-bound compounds that cause toxic effects as expressed by laboratory organisms; however these correlations are difficult to prove. These toxic effects may be due to synergism or possibly some compound(s) that were not analyzed in the first round of

chemical testing. The initial analyte list is, however, extensive and contains compounds expected to be present based on past discharge records. However if unusual bioassay results are observed, there will be further investigation to evaluate the potential causes.

Specific Comment 5:

Why is bacteriological testing not included? Bacterial, fungal and viral causes may be paramount in organism toxicity (and human health threat) independent of chemical toxin levels.

Response to Specific Comment 5:

Bacteriological testing is not routinely conducted in sediment studies, as the state-of-the art now stands. This comment raises an interesting point as to yet another possible contributor to observed toxicity. At this stage of investigation, i.e., characterization, bacteriological, fungal, or viral tests are not performed. After collection, evaluation, and discussion of the of the initial results with the agencies, possible contributors to any observed toxicity will be discussed and identification of sources will be assessed.