

Appendix G
Galveston Bay Indicators Project Final Report
Responses to Individual Reviewers' Comments

Contract Number 582-3-53330

Prepared By:
Geotechnology Research Institute
Houston Advanced Research Center
4800 Research Forest Drive
The Woodlands, Texas 77381

Principal Investigator
L. James Lester, PhD.

Project Manager
Lisa A. Gonzalez

Prepared for:
Galveston Bay Estuary Program
Texas Commission on Environmental Quality
711 West Bay Area Boulevard, Suite 210
Webster, Texas 77598

Contract Manager
Steven R. Johnston
Regional Monitoring Coordinator
Galveston Bay Estuary Program

May 2005

The authors of the Galveston Bay Indicators Project Final Report appreciate the reviews, comments and suggestions provided by scientists, stakeholders, and members of the public. Reviews of the methodology and final results of the Indicators Project were solicited in several ways, including an Indicators workshop held at the University of Houston on December 8, 2004; meetings of the Galveston Bay Council (the advisory body to the Galveston Bay Estuary Program) and its subcommittees; and individually submitted comments.

The methodology and results of the Indicators analyses were reviewed in great detail at the workshop and Galveston Bay Council meetings. In most instances meeting attendees reached a general consensus on a suggested change for an indicator. Feedback received at that workshop and Galveston Bay Council meetings was incorporated into the final report.

Comments and suggestions were also received from reviewers (generally agency and university experts) independent of the December 2004 workshop and Galveston Bay Council meetings. A compilation of individual comments received is included in the pages below as well as the responses by the authors of the final report (highlighted in red text). Names of reviewers who submitted the comments below have been removed, but their agency affiliation is provided. In all six individual sets of comments came from the following agencies and organizations:

1. NOAA Fisheries Galveston Laboratory
- 2a. Texas Parks and Wildlife Department, Coastal Fisheries Division, Dickinson Office
- 2b. Texas Parks and Wildlife Department, Coastal Fisheries Division, Dickinson Office
3. Environmental Institute of Houston at the University of Houston Clear Lake
4. Texas Commission on Environmental Quality Surface Water Quality Monitoring Program, Houston Regional Office
5. Texas Sea Grant/Texas Cooperative Extension

Comments were both editorial and substantive (i.e. regarding methodology and data) in nature. In some instances, a suggestion was immediately accepted and incorporated into the final report. In other cases the suggestions were given great consideration by the authors, but ultimately not integrated into the final report. Reasons for not including suggested changes included:

- Comments were contrary to feedback and consensus reached at the workshop or Galveston Bay Council meetings
- Suggestions were outside the scope of the Indicators Project
- General or philosophical disagreement on the part of the authors
- Change was not necessary or appropriate; in these cases clarification of the methodology was given by the authors
- Data availability issues

Again, the authors greatly appreciate all reviews and comments received. Even in instances where a suggestion was not integrated into the final product, the comment generated debate and a healthy discussion of the issue, indicator result, or underlying methodology. Peer review such as this can only strengthen the final results.

Thank you.

Jim Lester, Principle Investigator

Lisa Gonzalez, Project Manager

Note: Page numbers refer to draft version of the final report and do not match page numbers of the final version of the report.

Review #1 - Received from Scientist at the NOAA Fisheries Galveston Laboratory

Please take these as constructive suggestions and forgive me if I get a little testy (I should be able to edit those comments out, but one or two might sneak in). Some comments I'm sure you have already thought of but I include for my own knowledge. I will try to proceed through the doc page by page, and then save general comments to the end. I apologize for not having a more uniform style in my editing. It would have been nice to have this document in MS Word format. Maybe next time.

TOC – needs to update the Methods link to get the correct page number; probably you will do that for the entire TOC once finished. Are you going to add bookmarks or hyperlinks to these sections for easy navigation? **Done**

Give more complete titles to the appendices so the reader can find that batch of info easily; I know Appendix B deals with nutrients. **Done**

Page 7 paragraph 3 Line 6: “~~in order~~ to” **Done**

P7p3L5: “...the Bay...” no need to capitalize the Bay; everyone will know you are talking about Galveston Bay unless you say otherwise. This runs throughout the document, and looks haughty or privileged, or slang. **Kept “Bay” per other reviewers’ comments. “Bay” refers to a proper noun.**

P9p1L4: “..foundation” change to “basis” **Done**

P9p1L1: The indicators us a three-tiered approach.” might be more like: “To become an indicator, a parameter is assessed using a three-tiered approach.” **Done**

P10p3L3: “for the interpretation of the reader.” Might be more like: “for interpretation by the reader” **Done**

P10p5L2: “One indicator based on ...” might be: “The indicator that is based on...”, and later on: “Another indicator based on...” insert the “that is” before based on. **Done**

P10p5L6: explain what the 303(d) list means: you wait clear until the discussion to finally tell me that it is water not suitable for human recreation or some such. **Done ...the State of Texas List of Impaired Waters [the 303(d) list]**

P10: I think you might give a summary table for the 31 indicators evaluated with 3 factors: a status reflecting health of the bay as Poor, Average, Good; and a trend: Worsening, Stable, or Improving; and data (decades of): 1, 2, ...N

The primary purpose of this project was to develop the indicator framework. The list of indicators is provided in Appendix A. Status of specific indicators will be provided via Status and Trends and GBEP outreach publications.

Map of Galveston Bay.....add a distance scale (in miles) **Done**

P13p2L3: “Thus an indicator species...” might be “For example, an indicator species...” **Done**

P14p1L2 “A temperate estuary ~~will obviously~~ exhibits... **Done**

P14p3L2 “..a random fluctuation.”.....add “ ...fluctuation in the indicator.” **Done**

P14p3L5 “There are ~~two~~ three ways to solve...”

I think “The third way is to accept a standard screen that represents the optimal state for the indicator.” **Done**

P14lastpL2 “The GB system ~~is well~~ is monitored by public agencies ~~well planned sampling and assessment strategies~~ for the assessment and protection of its resources.” **Done**

P14lastp omit last sentence. **Done**

P15p2L7: spell out QA/QC **Done**

P15p2 omit last sentence **Kept original language**

P15lastpL2 “changes in ~~some~~ other...” **Done**

P15lastpL3 “a regulatory ~~and~~ or...” **Done**

P18p1L1 Why were the samples removed? Their values should have been set just below detection level if they could have been present, or set to zero if they might not have been present. But the samples were taken and analyzed. If there was another reason besides “below detectible limits”, that needs to be explained. **Nutrient samples listed by TCEQ as below detection limits were removed after a data quality review. The PIs thought it better to omit the data than to include those that may have been collected and designated as BDL because of sample error. Nutrients were treated differently than were other parameters such as toxics in sediment (BDLs retained). The PIs thought it more plausible to have toxics BDL than to have nutrients BDL, particularly in tributaries with known issues of eutrophication. However, per comments from several reviewers, the indicator was re-analyzed and BDL values were included.**

P18lastpL4&5 Why ere nitrate and nitrite combined into one parameter? Is that necessary? Some analytical problem or shortcut?

Nitrate, Nitrite, and Nitrate+Nitrite storet codes are combined to provide a complete nitrate-nitrite spatial temporal data record. This is standard procedure for the Galveston Bay Status and Trends Project which supplied the data. Additionally, all use one TCEQ screening level (see Table 1 in the final report).

P20 Table 3. I sure wonder about lumping these 4 factors into one. It would be easy to show all 4 in each cell instead of hiding it back in Appendix B. Maybe you could at least underline Appendix B in the text just before the table. Indicators are meant to provide a simple graphic for outreach to the public and decision makers. All four parameters are related to NPS pollution and can be shown as one indicator. Should the reader want more detail, he or she can refer to the Status and Trends project.

P21p1 omit last sentence. Kept original text

P21p3 Why the 85th percentile and not the 50th or 75th? explain with environmental relevance. The 85th percentile was chosen for two reasons: 1) it is a percentile utilized by the TCEQ for established screening levels, and 2) lesser percentile levels such as 50th and 75th percentiles (which were tested by the PIs) do not provide a fine enough scale for comparison.

P23p4L1 add title to Appendix C Done

P39lastpL3 “or 10 with 0) “Done

P41&42 The increase in average length hardly looks significant, and what about the CPUE? Is that statistically significant? Either way they should be explained. Was there no subbay area separation available? No appendix for these data separating out the 4 species? Would be nice to see. This is an indicator graphic. The detail described above (statistical significance and subbay distinctions) will be included in the Status and Trends Project.

P43p2L4&5 “...over the years in a decade, the effect of sampling effort will be limited.” Change to: “...decade, the variability due to changes in yearly sampling efforts will be reduced.” Done

P43lastp omit the last sentence -- it is not true, only if the former is the correct explanation. Done

P43lastp Need some statistical test here to support the conclusion of change. Indicators are meant to provide a simple graphic for outreach to the public and decision makers. Statistics are not included in the graphic. The PIs used statistical tests in the development of the indicator. Prior to decadal average calculations, a linear regression was run on the annual total species counts to determine direction and significance of annual trends: $R^2(\text{bag seine})=0.76, n=27$; $R^2(\text{gill net})=0.72, n=27$; $R^2(\text{shrimp trawl})=0.53, n=24$.

P44 Fig.6 change “taxonomic codes” to taxa , added “(numeric codes corresponding to taxonomic names)” in text and add “in Galveston Bay” to the title somewhere. Done

Regarding BIOL 5 & 6: Why was there no attempt to split into subbays like some of the earliest indicators? Were the data that poor? At least this desire to split should be mentioned in the Future Use section of each. Added to “Future Uses” Section, “This indicator was developed as a bay-wide indicator. Data are available and could be analyzed at the individual subbay level should the GBEP and its advisory subcommittees deem it appropriate.”

P49&50 Fig 8-11 Use same Y-axis scale and the same colors for the regions on all 4 figs – simplifies the comparisons. I object to 3-D charts because the reader can’t accurately interpret the value of the vertical bars; change the figs to 2-D if possible (and I know it is possible). In Fig. 10, where is the bar for East Bay for the 1980’s? Figures 9 and 10 (spring and fall salinities removed). Figures for winter and summer salinities converted into 2-D charts. All subsequent charts in the report are renumbered.

P50 Future Use Salinity is a measure of mixing of Gulf water, freshwater inflows, and local rainfall. Added

Also, I would not put emphasis on the change from 1990’s to 2000’s because you are dealing with half as much data in the 2000’s. The data are averaged to normalize for sample size.

P50 L3 change “fundamental” to “complex” Done

P51 Define “Low intensity and High intensity developed land” and “Bare/transitional land” Use examples of where these occur. What was the 9th category that was not used? Might list it too but in (). Added:

1. *Low intensity developed land* (Contains substantial amounts of constructed surface mixed with substantial amounts of vegetated surface. Small buildings, streets, roads, and cemeteries with associated grasses and trees typically fall into this subclass.)
2. *High intensity developed land* (Contains little or no vegetation, includes heavily built-up urban centers as well as large constructed surfaces in suburban and rural areas. Large buildings, interstate highways, and runways typically fall into this subclass.)
3. *Cultivated land* (Includes herbaceous (cropland) and woody (orchards, nurseries, vineyards, etc.) cultivated lands.)
4. *Grassland* (Managed and unmanaged grasslands; includes lawns, golf courses, forest or shrub areas converted to grassland, pastures, and areas of permanent grassland with altered species composition.)
5. *Woody land* (Includes non-agricultural trees and shrubs.)
6. *Woody wetland*
7. *Non-woody wetland*
8. *Bare / transitional land*

9. *Open water* (Not included in the indicator.)

P51-52 SOC/ECON1 needs more results. How much of each category is there now in each county, or the whole area analyzed? What are the units shown in the figure? If they don't make area-sense, i.e. are in 30x30m units, perhaps it would be better to cut that portion of the figure out.

Information added for each county. See indicator for revision.

P53 Census Why were earlier census data not included? 60's, 70's, 80's?

A line chart showing each county's population changes could be instructive; use log(base 10) scale for the Y-axis and decades along the X-axis .

While historical census data would be interesting, we wanted to keep the indicator graphic simple and chose to use the most recent US Census data. Historical population trends can be analyzed by the Status and Trends Project.

Future Use: Might start looking at ethnicity. (I see you brought this in the Discussion)

Ethnicity was analyzed, but was deleted from the report per instruction from participants at the Indicators Workshop in December 2005

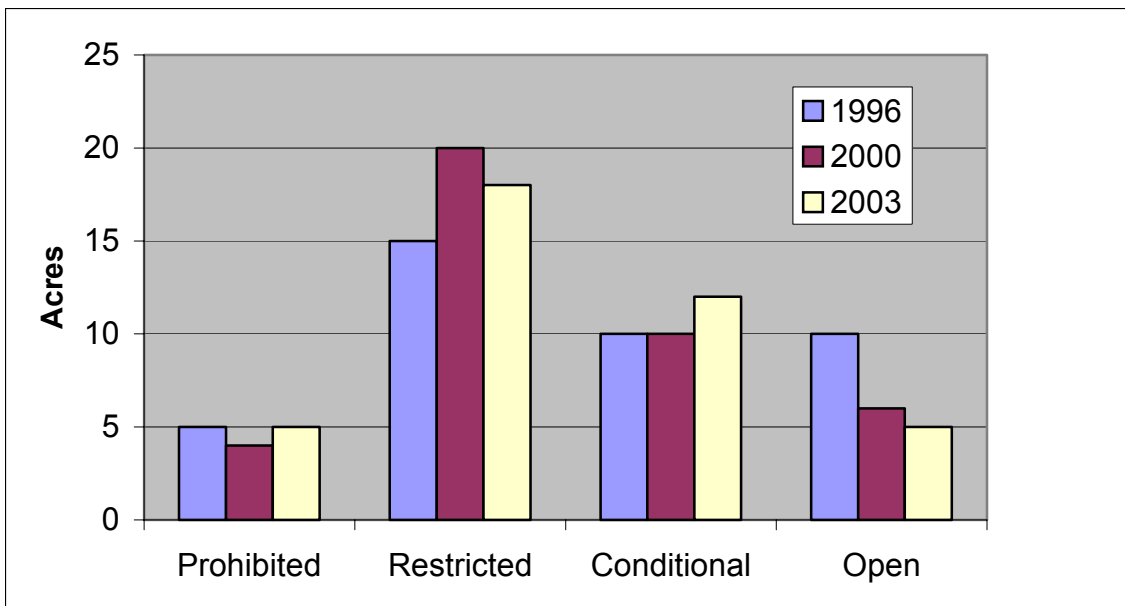
P55 Change surface area results to "increased". Recalculate areas and check the Fig 13 legend – the indicated areas are not the same. Recommend changing the speckled pattern to something easier to see, like red and white diagonal stripes.

This map is revised. ADV 20 sits on top of Adv 3 (in the upper reaches of the HSC) which remains in effect and did not change in area. Area did not change, but # of species and # of contaminants did.

P56 SEAF2 needs expansion. How about giving us the acres by decade by category?

We replaced the original chart (Figure 14) with a chart that follows the design below.

The new chart shows changes in acreage for each of the shellfish harvest classifications for the years 1996, 2000, and 2003.



Maybe also have 3 figures, one for each decade showing the spread of the conditions in G Bay.

Data are only available from the agency for the years 1996, 2000, and 2003. Cannot show change by decade.

Your indicator category “Condition Improved” might not be too good because it allows the change from Prohibited to Restricted to be equivalent to “Conditional to Open” or “Restricted to Conditional”. All those are not equivalent as far as oyster harvesting goes. If nothing else, bring this up as something to be looked at in the future.

This issue should be addressed by the above mentioned change to Figure 14.

P57 Fig 15. Retitle to change “Shellfish” to “Oyster and Clam” Kept original text. Shellfish is the term used by the DSHS.

P58 Give a list of the species used in a table.

List of 35 species analyzed in TPWD commercial landings data included in Appendix E.

P58 What does this mean: “prior to 1997 the data were not reliable”? Use the data and attach a qualifier if necessary. How are we going to see a change without it? The request for historic data was made. However, data from the TPWD were not available from the agency prior to 1997.

P59 Fig 16 Y-axis label “Weight (millions of pounds)” and adjust the scale numbers. Can’t even see Finfish.

Adjusted Y-axis to millions of pounds. Finfish are difficult to see-- their numbers range in the tens of thousands of pounds versus millions of pounds for other harvest classes. Steps will be taken to ensure proper formatting for this indicator when used in outreach materials.

If you really want the composite as the Indicator, just use a solid color for each bar and don’t split into groups. Participants at the December 2004 Indicators Workshop requested that the graphic be split into groups.

I want to see confidence limits on those bars. Confidence limits are not appropriate for a public outreach indicator graphic (refer to NEP Indicator reports of Puget Sound, Chesapeake Bay, and others). Confidence limits are appropriate for S&T products, but not in indicator graphics.

Fig. 17. Same idea: use solid bars without categories. Participants at the December 2004 Indicators Workshop requested that the graphic be split into categories.

I still think you should use the other data for the 70s, 80s, 90s to match Fig 16; add a qualifier if necessary. The request for historic license data was made. However, data prior to 1997 were not available from the TPWD.

Table 19: This might be nice to see in individual counties. It might show if the public is responding to Health Hazard Notices. Might add a part B to this by looking at it as a per capita index. If per capita use is decreasing, it is a sign that the bay is not functioning as well as it used to. **Included detail for the five counties. Per capita indicator is a good suggestion for a future indicator. Must be able to compare annual license data to decadal census data.**

Fig. 18: Why the artificial 10-yr periods? These do not correspond to the Fig 16 periods. I recommend changing to the 70s, 80s, 90s, and explain the limitations of the data. Another case of limited data? **Figure changed to standard decades.**

P60 FISH2: throw out the size—just too few data and too regulated by fishing laws to be meaningful. **Figure kept, but modified to show all species and standard decades.**

P62 Fig19: where is the 1974-82 data column?
Figure removed

P64p5 “Texas Senate Bill...” That paragraph is not needed because the same info is on p63. **Kept original text in case one indicator used without the other.**

P66 Fig 21: Need a more complete legend with matching colors to annual bars. Y-axis label might be “Millions of Acre-feet” and reduce the numbers accordingly.
The legend is included in Table 20. Y-axis scale changed to millions of acre-feet.

P66 Water 3 ...303(d) List, ie “impaired waters” explain here **added”... i.e. stream segments that do not meet water quality criteria established for designated uses.”**

P69 Last sentence, add: “volume of spills” **Done**

P71 Future Use: include a measure of the volume of each pumpout site or total pumped out. **Only data describing numbers of pumpout facilities and locations of facilities are available. Volume and # of uses are not recorded by the source agency (Texas SeaGrant).**

P72 Fig 23 Maybe add this in a per capita measure too; to see if this function is keeping pace with population increase. **Will consider this under future use section. Must decide how to apply annual data to decadal census data.**

P73 Are the recreational uses of the bay increasing or decreasing? Don’t you mean the recreational use of the bay? **Kept original text.**

P73p1 Delete next to last sentence. Add what types of data may be useful in last sentence, eg beach access points, or parks along the water. **Kept original text.**

P74 REC2 Why aren’t there earlier data? This is unacceptable. **The data are not available from the source agency prior to 1997.**

Again, should a part B be added by making this a per capita measure?

Will consider this under future use section. Must decide how to apply annual data to decadal census data.

P75 Future Use: Add that earlier years' data are needed. Added "Historical data should be provided by source agency for inclusion in the indicator."

P75 REC3: is useless. Why even include it? Detracts from the whole.

Participants at the December 2004 Indicators Workshop requested that this indicator be included.

P76 Value 1: lumping way too much nebulous data to be meaningful. Kept original text. Workshop requested that this indicator be included.

P76 Table 26: give values in "Millions of \$" Done

P78 Last sentence is inaccurate. There is no obvious increase. It is roughly stable. Text changed to: "Public attitudes favoring stricter construction limits in sensitive habitats appear to be on the increase from 1996 to 1998, but remain nearly stable from 1998 to 2000 as are attitudes towards increased funding for conservation efforts."

P80p3L11 "..., especially those" change to: "..., not only those" but also those with money values. Financial values are also difficult to verify and often consolidate parameters that we would like to separate and exclude. Done

P81 p1L1 Just the opposite is true or why would we be looking at them? Text changed to: "In some cases, e.g. salinity changes, a scoring system is not provided due to the lack of an established comparison value or screening level. In other cases, particularly with regards to indicators describing changing uses and user groups, it is the authors' belief that these indicators should be provided without a value judgment since values differ among individuals. For example, there are declining trends in the number of commercial fishing licenses and waterfowl hunting stamps. One can make the observation that the decline is occurring, but cannot place a value of "good" or "bad" on the trend since it depends upon the perspective of the reader."

P81 Water Sed. L2 From this and some ancillary data... Done

P82 Join the last 2 paragraphs. Done

P84p1L6 "Availability of such habitat may be in decline" change to "is" Done

P86 Explain the blue crab abundance due to fisheries independent data—is it stable? Or increasing? Text added:

However, the trends in blue crab abundance found in the fishery independent monitoring data do not match the pattern of the commercial decline. Blue crab populations sampled

by TPWD using shrimp trawls in Galveston Bay exhibit declining trends in Trinity Bay and West Bay only. Blue crabs sampled in all other Galveston Bay subbays using shrimp trawls exhibit no trends.

General comments>

Statistical analyses are not just for the elite anymore. They offer an important means of testing for differences and similarities. Any conclusions drawn about changes in indicator values must be supported by statistical testing. And if the data are inadequate to support this testing, such information must be supplied so that a push to acquire the relevant data can be initiated. Who is more qualified to reveal the database's strengths and weaknesses than the contractor who must use the data to develop indicators? No one. That is why I am dismayed that statistical tests have been left out, and comments on the database strengths and weaknesses have been left out. Without these two, all we have is a nebulous concoction of what might or might not be, not the foundation of "indicators" that can be held up for support of the bay's state of health.

To the dismay of some who participated in the review of the Galveston Bay indicator framework, some indicators are presented without supporting statistics. In some cases, statistics would be inappropriate because the indicators are based on exceedence of standards or on single measurements. Determining whether parameter levels exceed screening values does not require statistics. Indicators for nutrients, fecal coliform bacteria and seafood contamination are designed in this way. The issue is not a difference between distributions. It is whether a standard is violated. The statistical analyses go into determining what the standard should be (this is done by the agency responsible for setting the standard). This project does not include the development of screening levels. Other indicators, such as population change and land use change, are based on single measurements at a specified time and difference statistics can not be performed. In the case of indicators for biological resources, statistics have been performed during indicator development, but are not shown as part of the indicator graphic. Each of the species is scored based on the direction and significance of a regression analysis using the last 20 years of data. Only if the regression is statistically significant, is the species scored as increasing or declining. Non-significant trends are scored as stable despite the slope of the regression line. Reporting the results of the regression analysis did not add value to the statement that these are significant trends.

The Discussion contains too much restatement of what the indicators were based on, and the reader already has read that. What would be nice to see in the discussion is some comparison with other estuarine programs' indicators, or that there aren't any to be compared with. Probably Tampa Bay has something like our indicators. Have you looked there? What about the Corpus Christi Bay group? Do they have something comparable? Or the Puget Sound group?

A section on comparability to other estuary programs is included on page 75 of the final report.

I am quite disappointed in the lack of discussion on data problems. Many indicators are suffering from a lack of data, sometimes when there is data but it will not be released and sometimes when there should have been data but the responsible agencies did not collect it. Is this report another case of a state agency not being able to speak up for fear of hurting someone's feelings and fear of reprisal? How does this look when a contractor must take this stand? And how is the public suppose to shout for improvement if they are not given the information from which to decide future needs? It is a shame to miss this opportunity to provide potential direction to agencies as to what data are needed and desired. Maybe some statement could be added to the Discussion.

To the best of our ability these indicators match the properties defined in the introduction. All of the indicators described above are relevant to an assessment of the health of the bay or the sustainability of bay user groups. The conceptual soundness of these indicators has been validated by our method of obtaining potential indicators from credible sources and asking advisory experts to approve their inclusion. We have separated proposed indicators that proved not to be feasible at this time. They are discussed at the end of this section. In some cases, we had problems with the availability of data to complete an indicator and have described only the baseline or only some segments of the bay. Including the indicator in the above suite assumes the future availability of information to complete the assessment. Reliability is an unresolved issue for some indicators, not only those using data on biological resources. Some of the data collected, such as counts of nesting birds, have large experimenter errors associated with them. Other data, such as bag seine catch, have high inherent variability that may have nothing to do with responses to environmental stressors. Efforts to evaluate the reliability of these indicators should continue for some time. Financial values are also difficult to verify and often consolidate parameters that we would like to separate and exclude. Finally, we have fought the tendency to include more and more information and thus attempted to achieve a semblance of simplicity. Indicators are presented in simple charts when appropriate and, if a table is necessary, many are color-coded with scoring information for ease of interpretation.

Overall the indicators chosen are good ones and can be made to serve the Estuary Program well. However, the scaling of the indicators and the preliminary evaluations of the Bay conditions require additional review and discussion by Estuary Program staff and advisors.

Scale of the indicators was discussed extensively at the Indicators workshop and subsequent Galveston Bay Council meetings. Very often the scale (e.g. water quality in a tributary or subbay) follows the scale by which data are analyzed by the Galveston Bay Status and Trends Project. In other instances scale of an indicator was modified based on advice from workshop and meeting attendees.

- **The Texas Council on Environmental Quality (TCEQ) screening levels are just that, screening criteria used to identify problematic water segments. They are relative to the level of pollutants in other State waters. They have not been developed based upon known effects to fish, wildlife, ecosystem processes or suitability for human uses. This work proposes to use them in a manner to which they are poorly suited.**
- **The EPA will soon set nutrient criteria for this Nation's waters. The EPA effort should be acknowledged in the report and should play a role in the scaling of the nutrient indicators.**

While the TCEQ nutrient screening levels are not the ideal in terms of water quality standards, at the writing of this report they are the only nutrient water quality criteria available for use in the related indicator. The indicator can and should be revised to use the EPA or TCEQ nutrient criteria when they are implemented by the agencies. This is noted in the "Future Use" section of indicator, WSQ1.

- **Some of the discussion is subjective opinion (such as the comment on the Bay's natural turbidity) and should be supported or dropped.**

Language not changed.

When a water quality parameter exceeds a TCEQ screening level, it is acknowledged that some deleterious effects are occurring. This acknowledgement should spark some action in terms of seeking to reduce levels of that pollutant. Whether or not actions are practical or possible is not of concern for the purposes of this report. Superior ratings for areas that exceed screening levels on a regular basis (up to 10% of the samples) appear to be flawed. Most readers will think that a "superior" water quality level will correlate with "pristine" and not capable of improvement.

Based on feedback from the Indicators workshop and Galveston Bay Council meetings the scoring systems for many of the water quality indicators have been modified. The term "Superior" is no longer used and is replaced by the term "Very Good". In many

instances a parameter in a given area must have 0% of samples exceeding screening level or at the most 5% of samples with exceedences to achieve the score of “Very Good”.

In addition, the indicator scales used for the water and sediment quality parameters (nutrients, fecal coliform bacteria, etc.) do not always accurately portray what is known about water quality and its effects on fish, wildlife and human use of estuaries. This is because TCEQ’s water sampling methodology and screening levels are used to show trends, but aren’t necessarily designed to predict the effects of pollutants on bay health or suitability for human use. The bottom line is that TCEQ's screening levels for nutrients; when compared to what is known about the harmful effects of water pollution on aquatic biota or human use, maybe too high to prevent either damage to aquatic ecosystems or human health. Two examples of this can be seen when their screening criteria are compared to other’s efforts to set water quality standards for nutrients. Consider these examples:

1. TCEQ takes their fecal coliform water samples from the middle of a stream, one-meter down and only when it hasn’t been raining. This isn’t going to translate into the amounts of fecal coliform a swimmer or boater is going to be exposed to as they enter streams via the bank, ingest water from the surface and swim after it has rained). The TCEQ bacterial screening level itself appears to be too high to protect public health and recreational use of the Bay. Studies on this matter have been used to set fecal bacteria standards that are much lower than **400** colonies per sample. The EPA's water quality criteria use any single measure above **104** to indicate an unacceptable level of gastrointestinal diseases will occur and that average levels should be less than **35** (for enterococci in marine waters subject to swimming). Measures above these indicate that the waters are generally not safe for swimming. The current indicator uses the TCEQ screening level of 400 and this Indicators Report rates a water body that exceeds this level 10% of the time as "good". Swimming in such as waterbody would likely result in an unacceptable number of gastrointestinal illnesses. Showing such an area as a "green for go" color on a chart is not scientifically valid, keeps Galveston Bay's users uniformed about potentially dangerous swimming areas, and does not encourage users to clean up areas that need it.
2. The State of Florida’s proposed P standards vary from 40 to 10 parts per billion dependent upon whether the waterbody is fresh or saline and where it’s located. TCEQ’s screening level for P is 220 to 710 parts per billion.

Again, the TCEQ nutrient screening levels are not the ideal in terms of water quality standards. However, at the time of writing for this report the TCEQ screening levels are the only Texas-derived water quality criteria available for use with regards to nutrients and bacteria. Any disagreements with the TCEQ on how they arrived at their water quality screening levels must be taken up with that agency. The indicator can and should be revised to use the EPA or TCEQ nutrient criteria as soon they are

implemented by the agencies. This is noted in the “Future Use” section of indicator, WSQ1. It does not seem appropriate to apply water quality criteria from other states; this was the general consensus of the Galveston Bay Council.

Lastly, the discussion on Bay turbidity which states that Galveston Bay is shallow and has a mud bottom and is therefore going to be turbid, is a huge oversimplification of the system and dismisses a real problem that the Bay faces. Some Bay segments were deep and had a sandy bottom and were clear. For examples portions of upper Galveston Bay such as the San Jacinto River and Buffalo Bayou were probably clear water streams. The San Jacinto was by all accounts a tea-colored river with a sandy bottom. Its non-tidal reaches supported extensive clam beds that were recently (1970's) killed out by sedimentation. It had many side lakes with names like "Silver Lake". Its tidal portion consisted of sandy embayments with names like "Crystal Bay" and which supported submerged aquatic vegetation. Likewise, Clear Lake was probably clear at one time. Huge volumes of fine particle sediment of an anthropogenic source are entering the Bay and they are likely causing significant changes in its ecology (as well as its bathymetry). This is definitely an area where a knowledge gap exists (which could be easily filled by coring studies). The report shouldn't dismiss turbidity concerns out of hand.

- **The indicators are designed to show trends in bay health, therefore it is essential that bay conditions at the trend's baseline be accurately described. What does year zero represent? Was the monitoring begun when the bay was healthy or severely impaired? Does the trend line start from 100% health or 0% or somewhere in between? The report does not do enough to answer these questions. The report seems to use less than 10% exceedance of the TCEQ screening levels to establish its 100% healthy baseline. This is simply not supported scientifically. The GBEP needs to adopt of a science-based methodology for setting baseline conditions.**

It is critical that some information be added to the report that details the bay's state when monitoring began. For example in regards to fisheries, at the time monitoring began there was still a commercial fishery for redfish and speckled trout. People often ran large seines out into the bay and hauled in everything they could catch to the shore. Fishery populations may have been at a low historic level due to heavy exploitation. A steady or slightly increasing population trend from this point may not be a sign of increasing bay health but could actually signify release from over-harvest. Perhaps recovery of fishery populations should have been meteoric once management restraints were placed on over exploitation of the fishery?

Similarly, at the time monitoring of nesting colonies was begun, wading bird populations had yet to recover from their brush with extinction because of plume hunting when they were hit with the combined effects of DDT and fishermen who were stomping on the eggs in order to reduce their “competition” for fish. Therefore the modest recovery and declines since that time could be interpreted many different ways.

Water quality monitoring began just before modern sewage treatment became widespread; certainly waters were at their poorest historic quality. But what were conditions prior to European settlement of the region?

The Chesapeake Bay program uses the presettlement bay conditions as their 100% healthy condition. They've made an attempt to collect data on what the water quality, fishery populations, etc. was at that time. This is a valid way of setting baseline conditions and deserves consideration. It represents the conditions under which the ecosystem's organisms and biological interactions evolved. The GBEP should do the same.

While the authors agree that it would be ideal to assess indicators against historical baselines of “presettlement bay conditions”, this is simply not feasible for the Galveston Bay Estuary given that the data are not available. Indicators assessed against such a historical baseline would be based largely on assumptions or anecdotal information. The major data sets describing Galveston Bay were begun by regulatory agencies after implementation of landmark environmental legislation in the late 1960s and early 1970s. The indicators can only utilize the data that are available.

This discussion on turbidity relates to the use of turbidity as an indicator in national indicator programs. To use turbidity as an indicator to compare Galveston Bay to a bay system in Florida or elsewhere along the East Coast of the United States is not appropriate.

The wading bird guild (Great Blue Heron, Tri-colored Heron, and Snowy Egret) share a common strategy when foraging for their young, that is they tend to congregate in drying fresh and brackish wetlands for fish and invertebrates made available by the shrinking pools of water. These wetland types have shown the greatest losses and continue to be lost at a high rate, yet the report only discusses changes in fish numbers or nesting islands as a possible cause for population declines.

In addition to the above comment, trading out the snowy egret with the great egret, or perhaps simply adding the great egret may make changes in environmental conditions more apparent and should be discussed.

Indicator Biol 4 specifically addresses colonial nesting bird population trends. Based on suggestions received at the Galveston Bay Council meeting, the white ibis was added as an indicator species to specifically address the declining wetland habitat issue. Neither this species nor the great egret show significant declining trends for the period 1973-2003.

Editorial Comments:

Acronym use versus use of full titles is inconsistent in table titles.

Inconsistencies were corrected.

Consider using a color scheme that will be apparent in black and white copies.

Attempts were made to add hatch marks to the graphics; they did not look good. Any outreach materials developed using the indicators will keep the black and white graphics issue in mind.

Review # 2b - Received from Scientists at the Texas Parks and Wildlife Department, Coastal Fisheries Division, Dickinson Office

Page 58-59: Regarding the discussion comparing declines in commercial landings with declines in commercial licenses, it should be noted in the report that this time frame coincides with implementation of various commercial license management programs in Texas where commercial licenses (bay shrimp, bait shrimp, crab, and finfish) are permanently retired from the fishery through a voluntary buy-back program. The reduction in licenses as depicted in Fig. 17 is a result of buybacks and natural attrition and are not being replaced in the fishery. This data is available from TPWD for Galveston Bay, by fishery, by year.

Language describing the voluntary buy-back program was added to the indicator, Fish1, just after Figure 15.

Page 59-62: I don't believe the number of fish landed by recreational fishermen is an appropriate indicator of relative health of the fishery in Galveston Bay. There are several components that can affect landings that may have no affect on bay health. For example, changes in angler attitudes and behavior can affect landings. Survey data following the same group of fishermen over the last decade indicates there has been a shift toward the practice of catch-and-release among saltwater anglers in Texas. Also, there should be some mention of the effect more restrictive management measure have on landings. For example, the landings depicted for the '74-'84 decade in Fig. 18 covers the time frame where there were very liberal or no limits on recreational harvest. Bag and size limits began to be utilized in the mid-1980's. No significant change in these limits has occurred since 1995 (except for flounder).

In addition to developing ecological indicators that describe the relative health of the bay, the Indicators project also developed indicators based on Galveston Bay user groups, recreational fishermen being one. Language describing catch and release and restrictive management regulations was added to the indicator, Fish2.

I'm not sure of the significance of Fig. 19 as it relates to indicators of species health in Galveston Bay. There is no explanation of the species used in developing this graph. The same considerations regarding major regulation changes and attitudinal/behavioral changes by fishers apply.

The original Figure 19 (Average length of recreational landings of fish captured in Galveston Bay by decade, 1974-1998) was removed based on this comment and feedback received at the Galveston Bay Indicators Workshop.

Page 70-71: Future reports should include the number of US Coast Guard documented vessels, especially when evaluating pump-out facilities as vessels of this size class are more likely to utilize these facilities. Only within the last year have USCG documented vessels (recreational and commercial fishing) been required to also obtain registration through TPWD.

Noted.

General comments

It would be helpful to incorporate something other than color to identify the different ranking levels if the document is to be made available electronically.

Attempts were made to add hatch marks to the graphics; they did not look good. Any outreach materials developed using the indicators will keep the black and white graphics issue in mind.

Review # 3 - Received from Scientist at the Environmental Institute of Houston at the University of Houston Clear Lake

Thank you for the opportunity to review this report. I apologize for these late and brief comments, so I have tried to focus on the more major issues. I believe that this represents a significant contribution toward the knowledge and management of Galveston Bay. In addition, this was obviously a very hard task to accomplish, especially for some categories of indicators. The challenge of combing and analyzing diverse environmental data is daunting for anyone. This report helps illustrate where data gaps occur and should provide a common point of reference for future analyses. I also apologize for any comments that should have been forwarded at an earlier review or workshop prior to my arrival back to UHCL. Obviously comments regarding selection of indicators would have been more useful at that time. These comments should probably be reserved for any future assessments.

I do have some general and specific comments and suggestions regarding this report. They are listed below.

General Comments.

WSQ1. The ranking scale deployed by this investigation is in my opinion problematic. It is not “symmetrical”. Obviously, the use of and qualitative ranking scale will be considered biased by many, based on the view of each observer, and their political and social perceptions. However, the use of a more symmetrical ranking would go far to remove many of these criticisms. For example, one such scale could be 0-10% (Superior), 10-25% (Good), 25-75% (Concern), 75-100% (High Concern). This would provide a more balanced evaluation of when a particular segment is “unusual” in regards to coast wide screening levels or legal standards. The authors are correct in stating that perhaps different, more conservative approaches should be followed in regards to toxic chemicals. However, a more symmetrical approach should also be employed in this case. **The ranking scale has gone through several iterations, the first of which was very similar to the scale described above by the reviewer. In the end, the PIs and workshop reviewers decided upon the more conservative scale used in the final report.**

WSQ1. In some cases the authors fail to clearly define the importance or lack of current understanding of the level of particular variables. For example, TOC is identified as being below or above the 85th percentile. There should have been more discussion about the implications of low TOC and/or other nutrient values. Perhaps in some cases these low values may also represent some kind of “level of concern” or limiting condition that could detrimentally limit natural productivity. In this case they should be also be flagged as appropriate. Again, the response of living organisms to these variables is more appropriate described as unimodal versus linear.

The PIs initially considered employing a two-threshold system with a lower threshold limit and an upper threshold limit. However, since the upper threshold limit was determined by state screening levels, in the end, the PIs felt setting a lower threshold

would be too subjective. Will suggest development of lower threshold for the nutrient and TOC indicators in “future use” section.

WSQ1. In other cases it is unclear whether screening levels were actually calculated from values from the universe of Galveston Bay segments or from coast-wide data as done by TCEQ. This should be clearly defined in all cases.

The screening level for Total Organic Carbon (TOC) was calculated using data from the Galveston Bay Status and Trends database (TCEQ screening level not available). The TOC indicator was removed.

A following more serious concern is the handling of non-detects in construction of percentiles and determination of rankings. During several evaluations, values below detection limits were not used. To me this is a serious issue that needs to be carefully re-examined in the evaluation. Removing non-detect data from the analysis is paramount to throwing away information, and more seriously, in some cases can skew the distribution of sample statistics and percentile values.

Samples listed by TCEQ as below detection limits were originally removed after a data quality review. The PIs thought it better to omit the data than to include those that may have been collected and designated as BDL because of sample error. However, per comments from several reviewers, the indicator was re-analyzed. A number of nutrient samples were reported by the agencies as being below detection limits. Data with below detection limit values were set to zero and included in the analysis.

WSQ1. Any particular reason air and water temperature, turbidity, orthophosphate, dissolved oxygen and ammonia nitrogen were not evaluated?

Since these are common parameters that are measured and/or can be combined with other nitrogen measurements (= total N) some discussion should be included as to why they were not analyzed. Historically the HSC and many tidal bayous (e.g. Dickinson, Buffalo Bayou) have exhibited low dissolved oxygen values. In addition, HSC has exhibited elevated ammonia readings in the past. A lot of data exists, in some cases continuous data (e.g. open bay TWDB), to make long term comparisons. In the case of dissolved oxygen, data could be sorted by time period to insure comparing only values from morning readings if needed. Also, the concern regarding global warming and continued use of power plants should warrant monitoring at least at the local scale of air and water temperature. In addition, these variables are drivers for other chemical reactions.

- Water temperature data are routinely reviewed by the Galveston Bay Status and Trends Project. Air temperature data can be added to Status and Trends analyses if suggested by the Monitoring and Research Subcommittee.
- Total phosphorus was analyzed because the TCEQ data record for this parameter is more complete than for orthophosphate.
- Turbidity (secchi disk) data are not analyzed by the Status and Trends project due to the subjective nature of the measurement. Instead, the Status and Trends project uses Total Suspended Solids as the turbidity parameter of choice. TSS data tend to highly variable and are influenced by weather-related variables. Given this, the PIs felt that TSS was not suitable as an indicator of Bay health.

- Ammonia nitrogen is included in the nutrient indicator
- Dissolved oxygen data are problematic in that they are collected at all times of day. Dissolved oxygen problems in Galveston Bay tributaries are related to eutrophic conditions. Eutrophication is addressed in the nutrient indicator.

WSQ general. I realize that environmental data is notoriously noisy, but the use of average values can be misleading. I strongly recommend that error bars (+/- 2SE) be placed around estimators that employ average values in their ranking.

The indicators will ultimately be used for public dissemination. Error bars are too technical for public outreach materials.

WSQ general. One major oversight. The authors should include point source loading (e.g. TSS, BOD, Flow, ammonia, and/or metals & organics) in their indicators and some index of environmental enforcement/non compliance, per violations. This should be added.

Added to the list of future indicators in the Discussion section.

WSQ general. I understand the focus on using government databases due to their spatial and temporal coverage. However, for sediments two large databases exist (EMAP and REMAP, pre NCA) and BEG. These data sets should have been included in the data analysis.

Will be recommended for future use.

WSQ general. Why delete below detection limit values from percentile computations? No explanation given. See general comment 3. Also, does TCEQ delete below detection limit values for computation of their “screening levels” which I believe are statistically derived from the only from ambient estuarine samples collected from 1990 to 2003??

See response to number 3, above.

WSQ1. How about low nutrient levels? Can that be a concern? Some discussion should be included.

That text is included in paragraph 1 of Indicator, WSQ1 under the section, “Relation to Assessment Question”.

WSQ general. Should mention, that TOC in some cases can be used as a surrogate of BOD measurements. Also, you mention the use of TOC and chlorophyll-a together, but fail to explain why. For example is high TOC and chlorophyll-a mean something?

TOC removed as an indicator per other reviewers’ comments.

WSQ general. Deletion of values below detection limit. See comments in general section 3.

See response to number 3, above.

WSQ general. Rating scale too coarse and unsymmetrical, very conservative, risk adverse. Should consider adding additional categories, superior (0), good (1-10), fair (11-25) (poor 25-75), extremely poor (>75).

The ranking scale has gone through several iterations,. In the end, the PIs and workshop reviewers decided upon the more conservative scale used in the final report.

Again, the inclusion of EMAP, REMAP and BEG data might expand the data set and make it possible to make more comparisons where n is small.

Text added to future uses section.

WSQ general. Assigning values below detection limits a value of zero should not be done without considering other alternative scenarios. See comments in general section 3. If this is done, you should also run with $<dl$ values = dl and/or $\frac{1}{2} dl$, or use some statistical technique designed to handle below dl values, to generate summary statistics (Helsel 2005).

BDL values assigned a value of zero based on TCEQ advice to Galveston Bay Status and Trends Project in 2003.

WSQ3. What is a “consultation”? Is this a study or something?

Consultation is the term used by the Texas Department of State Health Services to describe their tissue contaminant studies.

WSQ3. Rating scale too coarse and unsymmetrical, very conservative, risk adverse. Should consider adding additional categories.

The ranking scale has gone through several iterations,. In the end, the PIs and workshop reviewers decided upon the more conservative scale used in the final report.

WSQ4. Might also want to include the 30 day average value of 200 cfu/100 ml as another, more conservative benchmark or at least discuss not using it.

Text added to future use section.

WSQ4. Good point about the use of E. coli vs. Fecals etc in future. Question, how are you going to track change into the future when you switch indicators???

This is an issue that continually occurs when an agency changes sampling methodologies. When the change occurs, GBEP must work with its technical subcommittees and agency personnel to ensure that it is done properly. Text has been added to the future uses section.

BIOL2. This indicator is deceptively simple. However a functional salt marsh is difficult to define. Future indicator development needs to include information regarding stem density, acreage, some multiplier of the two (i.e. production on annual basis) and utilization by marsh organisms. Albeit, the last metric would be difficult to define.

The data above are currently not available. Additional text describing data needs was added to future uses section.

BIOL3. Need to point out that fishing pressure can also affect oyster numbers, not just environmental conditions (see page 40 predatory fish discussion). Also, TPWD used to monitor oyster landings and conduct limited oyster boat counts.

Text added to Relation to Assessment Question section: “Abundance is dependent on fishing pressure, habitat quality, and environmental stressors.”

BIOL4. No comments, area outside expertise.

BIOL5. good discussion of predatory species. However, should also point out these upward trends could also be caused by more restrictive fishing regulations and stocking programs, both of which have occurred over the last 30 years. Also, remember the baseline used for this assessment is the 1970’s a period in which commercial fishing was high for these species, pollution was severe and few commercial or recreational limits were present. In other words we may still be far away from the ideal carrying capacity for this system. Early writings (pre-1940’s) document a much more productive system with major fisheries that have declined. Also, tarpon, a major top level predator, which used to be extremely abundant in the 1950-60’s, is extremely rare. No satisfactory explanation has been given.

Text added to 1st paragraph: “However, it must also be noted that improving trends could also be caused by more restrictive fishing regulations and stocking programs, both of which have occurred over the last 30 years. The oldest historical data available as a baseline for this assessment was collected in the 1970s, a period in which commercial fishing pressures were high, few commercial or recreational fishing regulations were implemented, and pollution was severe. As a result, the indicator may show improving or stable trends, when in actuality; Galveston Bay may still be far from its ideal carrying capacity. Early writings (pre-1940s) document a much more productive system with major fisheries that have declined and species that are now rarely found in Galveston Bay (e.g. tarpon).”

BIOL6. Another item that needs to be discussed is the composition of this species diversity. Major changes in species composition have occurred over various time periods, for example during the early 1980’s few gerreids (mojarras) were collected, and in recent years more where. Changes in the composition are not captured by simple diversity indices (i.e. species count). These changes may be due to changing salinity regimes, temperature, improving water quality etc.

Species composition tracking will be addressed by the Status and Trends Project.

PHYS1. Local rainfall patterns within the watershed could also be used to supplement Trinity River flows. Trinity River flows are not good indicators for West Bay and Christmas Bay. Also, other gage stations, although not possessing as long a time series, could be used to supplement the Trinity River gage.

This issue has been examined. Gage data with the desired temporal and spatial coverage in other parts of Galveston Bay are not available. The indicators chosen use two approaches given the data available: USGS gaged flows and TWDB modeled flows.

The authors make an excellent recommendation, that is, the need for more long term gage sites in major rivers and local watersheds. This is a major problem in the Galveston Bay monitoring network.

PHYS2. Any particular reason why continuous salinity monitoring data was not evaluated? Unless it has been discontinued, the TPWD has collected, under contract to TWDB, a lot of historical data for open bay locations. Given the interest in freshwater inflows, these data sets should be examined.

Thank you for the suggestion. The PIs will look into acquiring the continuous monitoring data through the Status and Trends Project.

PHYS2. Also, why was TPWD salinity data not used as well? They collect salinity data during bag seining, trawling, gillnetting, and oyster dredge sampling. Obviously, these are collected at different time scale and using different methods, but they can be used to supplement TCEQ data, that in some cases, has very little spatial coverage. Also, TCEQ almost always uses specific conductance data. Given the choice, I would have used that data and converted to salinity using known published conversions factors as needed. It is not that complex an exercise. A spreadsheet conversion is all that is needed.

TCEQ specific conductance data is being considered for inclusion in the Status and Trends analyses. If a successful data conversion (specific conductance to salinity) can be completed, then it could be used in implementation of the indicator. This suggestion has been added to the future use section with the following caveat. Specific conductance is a better measure of low concentrations of ions in freshwater and salinity is preferred as a measure of the salt concentration of seawater. The possibility of combining salinity and specific conductance to obtain an indicator that could be used in the tributaries as well as the Bay was rejected. Combining the two parameters is complex and can introduce error.

Page 53-54. Seafood advisories. Some mention should be made that some of these trends may do to increased surveillance activities, not that more organisms have become contaminated (i.e. we are now looking harder). I am confident that we probably had more contaminated seafood in the past (1960's and 1970s) but we were not looking for them or not monitoring at low enough levels.

This is an indicator for which the baseline or historical data are not available. The authors chose to frame the discussion of consumption advisories around the available data (1990 to present).

Page 55. Shellfish harvest areas. Should mention that one of the major drivers for this indicator is microbiological contamination and the continued consumption of raw shellfish. Also, there will probably always be closed waters due to the presence of some background levels of indicator bacteria due to high warm-blooded animal populations, bypasses, and improperly designed onsite wastewater systems.

Language was added to the 2nd paragraph of SEAF 2: "The DSHS periodically modifies the status of shellfish harvest areas along the Texas coast as conditions warrant. The principal monitoring parameters are the fecal coliform concentration of the waters over the shellfish and the proximity to wastewater sources. Bacteriological conditions in the shellfish areas typically drive shellfish harvest area changes. However, other influencing factors can include rainfall, runoff, flooding, hurricanes, other extreme weather conditions, major spills, algal blooms, or the operation of wastewater treatment facilities. It is likely that some areas will always be prohibited as a result of high background levels

of indicator bacteria from local animal husbandry operations or wastewater treatment systems.”

Page 57 – Data reliability?? I assume this is effort data. Care to expand on this? That is a lot of pre-1997 data constitutes up to 30 years of information. If that data is not reliable I would hate to break that news to any elected official who provides funding to TPWD.

This is not the fishing effort data. This is commercial license data for which data are only available for 1997 to present. The language was modified to read: “Commercial fishing boat data were obtained from the TPWD for the years 1997-2003. Data prior to 1997 were not available from the agency. Therefore, it was not possible to complete a historical comparison of fishing boat licenses and commercial landings dating back to the 1970s.”

Why remove the data describing the ex-vessel values from the analysis??

Economic value data were removed because an indicator of sustainability would require the annual values for commercial and recreational fishing to be normalized to some set of constant economic conditions that could not be defined by the group of participating experts. Additional comparisons of weight and ex-vessel value will be made available via the Galveston Bay Status and Trends Project.

The authors are commended for suggesting that commercial fishing license data should be obtained for future comparisons. Also, they should strongly point out that without effort data (# vessels, tows etc), you cannot calculate CPUE and cannot adequately address questions involving density, overcapitalization etc.

Page 58-59. Recreational fishing

Did not TPWD provide creel data for analysis?? If I am not mistaken, I believe that part of their recreational creel survey includes information on where people come from (fishing effort). It seems that that data could be used to normalize fishing effort based on recreational licenses sold in the surrounding counties. Somewhere in this discussion, a comment needs to be made regarding the decline of average sizes, and changes in size and bag limits. Is there a relationship??

The suggestion to normalize the recreational effort using license data and the creel census data has been noted for the Status and Trends Project. Language regarding more restrictive management regulations was added.

Page 63. Annual gauged freshwater flows.

Define Max H, Min Q and Min Q-sal. I don't think these terms have been previously defined.

This indicator has been modified substantially. Of the three target flow volumes, only Min Q is used in the Indicator. It is described in full on page 51 of the final report.

Page 66. The authors should mention that 303d listing is partially a function of additional surveillance and changing water quality standards.

Language not changed.

Page 68. May want to calculate a percentage of vessels spilling. Normalize for less vessel traffic. Also, the *oil spill intelligence report* and USGS central probably has spill data going back into at least the early 1990's.

The Status and Trends Project will look into availability of data via the *oil spill intelligence report* and USGS central. The percentage of vessel spills in relation to vessels is small and doesn't add additional detail. Volume of vessel spills was added based on comments from Indicators workshop and Galveston Bay Council.

Page 69-78 – No comment, not reviewed.

Page 80. The authors suggest that Galveston Bay is not eutrophic, in contrast to EPA. Even though nutrient levels and chlorophyll levels are declining there are still areas where we have massive algal blooms with large fluctuations in dissolved oxygen and fish kills. I do believe the jury is still out on this subject, due to our poor understanding of the dynamics of nutrients in estuaries. I believe the high turbidity as expressed by TSS or Secchi disk readings, which by the way were not reported on and should be since they were referenced, have reduced the likelihood of experiencing some of the negative effects associated with nutrients. In other words during many times of the year, it is light not nutrient limitation that controls algal populations in Galveston Bay. A common progression of events I have seen is increased rainfall, highly turbid tributary, deposition of fine particulates and clearing of water, and subsequent increase in algal population or bloom.

Noted.

Page 81. I think the authors should strongly point out that there is a need to collect more fish tissue data in Galveston Bay. I believe this is a major limitation on understanding the flow of contaminants through the system.

Comments were not added given the DSHS plans to increase tissue sampling throughout Galveston Bay over the next few years.

Page 83. Second paragraph. Good point regarding the relationship of fishery regulations and fish size. This could effect any interpretation regarding environmental effects.

Page 83. I disagree; a diversity algorithm would have helped interpret this. That is a few rare species showing up can greatly influence a decadal index based only on species number, whereas it would hardly register using a diversity index.

Language not changed.

Page 84. I believe the authors should have examined the use of rainfall gage data as a surrogate for local watershed level contributions to freshwater inflow.

Noted.

Page 85.

I must have missed something, because I see no presentation of yield per fisherman (CPUE) data discussed in the text. The only discussions I saw were landings (per decade) and number of licenses sold (per decade).

Language not changed.

I do not believe you can conclude based on the data presented, that the reason for increased landings is due to increases in population size. Could be due to more effort.

Noted.

Good discussion regarding the average size of predatory fish and fishing size limits.

Pages 88-90. Very good recommendations for inclusion of new indicators, especially air stuff.

One major oversight. The authors should include point source (NPDES and State permitted) loading in their indicators and some index of environmental enforcement/compliance or lack thereof. If you are tracking spills you should also track point source loads which are much more abundant.

Noted.

Also, more monitoring of tributaries for biological communities is needed.

Noted.

Review # 4 - Received from Scientist at Texas Commission on Environmental Quality Surface Water Quality Monitoring Program, Houston Regional Office

I have 2 "big" comments on the draft. One is about the not using nondetects. Do you mean that you didn't use a value for those or that you removed them from the database? If you deleted them entirely, then it seems that the results would be skewed high, especially for parameters that have a lot of non detects like organics in sediments. If you didn't take them out of the total you divided by, you should explain it better, because the text seems to indicate that you did.

Samples listed by TCEQ as below detection limits were originally removed after a data quality review. The PIs thought it better to omit the data than to include those that may have been collected and designated as BDL because of sample error. However, per comments from several reviewers, the indicator was re-analyzed. Data with below detection limit values were set to zero and included in the analysis.

The other comment is that I don't think you should use WSQ-2 (TOC concentrations) at all as an indicator. No one even knows if it should be high or low. You treated it as though it should be low, but we don't know. I think that it should be left out entirely, unless there is some information somewhere that shows what the numbers are supposed to be in a healthy ecosystem.

The TOC Indicator was removed based on this comment and comments received at the Indicators workshop and Galveston Bay Council meetings.

Review # 5 - Received from Scientist at Texas Sea Grant/Texas Cooperative Extension

In terms of impervious surfaces--the number of watersheds with low imperviousness would be very interesting. The total picture is of course important, and perhaps the easiest indicator for a watershed-wide evaluation, but we really want to know how many of the subwatersheds are being impacted--are we paving over presently low-impacted watersheds, for example?

The indicator, Soc/Econ 1, deals with change in land use patterns: developed and undeveloped land. Spatial scale is always an issue. This indicator looks at developed land rather than imperviousness (which was originally proposed) at the county level. A subwatershed analysis should be proposed to the Status and Trends Project.

I have pretty much a similar comment with respect to biological resources--the amount of habitat remaining. Its not just the total amount remaining, but its connectivity and fragmented nature that is critical as well. But don't know if there is a simple measure out there for that.

An assessment of habitat fragmentation is under development by the Status and Trends Project.

Looks like a really great job overall! Congratulations!!

In some recent discussions and as the result of a recent paper, some of us are leaning towards "developed area" as a better indicator than imperviousness. I think I have been using imperviousness as a fancy way to say developed area, but some areas are more impervious than others, and so the term can get bogged down. Developed area might be a cleaner way to address the issue--and it lines up with the fragmentation concepts better as well.

Soc/Econ 1 was changed from "Change in land use patterns and impervious cover" to "Change in land use patterns: developed and undeveloped land".