SANTA ANA RIVER DISCHARGERS ASSOCIATION (SARDA)

December 21, 2018 submitted to: commentletters@waterboards.ca.gov

Jeanine Townsend, Clerk to the Board
State Water Resources Control Board
PO Box 100
Sacramento, CA 95814-2000

Re: Comment Letter - Toxicity Provisions

Dear Ms. Townsend and Members of the Board:

On behalf of the Santa Ana River Dischargers Association (SARDA), we submit this comment letter regarding the Proposed Establishment of New Toxicity Provisions in the Water Quality Control Plan for Inland Surface Waters, Enclosed Bays, and Estuaries of California.

The SARDA agencies all produce high quality recycled water and have a long history of consistently passing monthly toxicity tests. Our foremost concern is that the proposed TST procedure presumes our treated wastewater is "toxic" despite more than 20 years of test data showing otherwise. If the State Board adopts this presumption as their official position, it will undercut decades of hard work convincing the public to accept recycled water as "clean and safe." At a time when the State Board is asking everyone to make greater use of recycled water, in order to cope with drought and climate change, this policy makes it harder to achieve the goals of the State. Additional comments are attached.

If you have any questions, please call me at (951) 928-3777 extension 6327 or email me at javiera@emwd.org.

Sincerely,

Alfred Javier
SARDA Chairperson

Encl.: 14 page comment letter attached.

Comment #1: The Test for Significant Toxicity (TST) cannot be used to certify compliance with an effluent limit for whole effluent toxicity in an NPDES permit because it is not part of the federally-promulgated method and has not yet been approved by U.S. EPA as an Alternative Test Procedure (ATP).

1.1) Where EPA has established a standard test method, in accordance with 40 CFR Part 136, federal regulations mandate that dischargers must use these water quality monitoring methods to demonstrate compliance with effluent limitations or other conditions specified in a NPDES permit.[1]

> *"If EPA has 'approved' (i.e. promulgated through rulemaking) standardized test procedures for a given pollutant, the NPDES permitting authority must specify one of the approved testing procedures or an EPA-approved alternate test procedure for the measurements required under the permit."*[2]

1.2) Whole Effluent Toxicity ("WET") test procedures were promulgated and approved as standard test methods by EPA in 2002.[3] The actual test procedures are described in a series of method manuals.[4] These manuals, and the specific procedures for each WET test method within each manual, are specified at 40 CFR Part 136.3.[5]

1.3) The Test for Significant Toxicity (TST) is not discussed or described in any of the WET test manuals that were published as part of the methods promulgated in 2002. Rather, the TST procedure first appeared in a non-binding guidance document released eight years later.[6] To date, EPA has not promulgated the TST statistical approach under 40 CFR Part 136 or approved it as an Alternate Test Procedure (ATP).

1.4) Federal regulations prohibit any modification of an EPA-approved Clean Water Act analytical method for method-defined analytes.[7] According to EPA, *"method-defined analyte means an analyte defined solely by the method used to determine the analyte."*[8] And, the *"determinative technique means the way in which an analyte is identified and quantified."*[9]

[1] 40 CFR 122.44(i); 40 CFR 122.41(j)(4); 40 CFR 122.21(j)(5)(viii)
[2] 67 FR 223, 69952 (Nov. 19, 2002) [emphasis added].
[3] 67 FR 223, 69971 (Nov. 19, 2002)
[4] See, for example, Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms. EPA-821-R-02-013 - Fourth Ed., October, 2002.
[5] See Table 1A: List of Approved Biological Methods for Wastewater and Sewage Sludge
[6] U.S. EPA. National Pollutant Discharge Elimination System Test for Significant Toxicity Technical Document. EPA-833-R-10-004. June, 2010.
[7] 40 CFR 136.6(b)(3)
[8] 40 CFR 136.6(a)(5)
[9] 40 CFR 136.6(a)(3) [emphasis added].

1.5) Whole Effluent Toxicity is a "method-defined analyte." [10] A method-defined analyte is one *"that does not have a specific, known composition and the analytical result is dependent on the measurement technique. As a result, a change in the analytical technique has the potential to change the numerical value of the sample result."*[11] Such modifications can only be authorized for use in NPDES permitting through a formal rulemaking process like the one used to promulgate the original WET test methods.[12] EPA has declared that:

> *"A proposed test procedure will be considered a new method if it employs a test species, an endpoint or organism response, or a toxicity test concept that is not represented in the battery of Agency-approved WET methods. Since WET is a method-defined analyte, EPA generally considers the use of new test species, endpoints, or test concepts to be substantial changes, and therefore will be approved as new methods… EPA expects that Alternate Test Procedures (ATP) may include, but are not limited to, changes to the following aspects of an approved WET method: … test concentrations, dilution factor, or number of replicates …[and] method of data analysis."*[13]

1.6) EPA's Technical Document explicitly admits that the TST *"is an alternative statistical approach for analyzing and interpreting valid WET data."*[14] The TST changes the "method of data analysis" because it: 1) reverses the traditional null hypothesis, 2) introduces a new test concept called the Regulatory Management Decision (RMD) threshold, 3) compares control data to only one effluent concentration rather than to five different effluent concentrations as required for all WET tests used in NPDES permitting, 4) relies on new statistical procedures (e.g. Welch's T-test) not previously described in the WET method manuals, and 5) recommends increasing the minimum number of replicates to reduce the risk of error when using the TST to make compliance determinations. Consequently, the TST is clearly a new and different *"determinative technique for the way in which the analyte toxicity is identified and quantified."*

For method-defined analytes, the analytical techniques used to determine the presence or absence of toxicity is part of the method; therefore, any change to these procedures constitutes an impermissible modification to the approved method. In fact, when the Virginia Department of Environmental Quality considered regulating whole effluent toxicity using a Percent Effect approach that was virtually identical to the RMD threshold now proposed in the TST, EPA informed the state that such a change could not be applied in NPDES permits until an ATP was approved. Thus, an ATP is also required before the TST can be used to determine compliance in lieu of the promulgated WET test procedures.

---

[10] 67 FR 223, 69966 (Nov. 19, 2002) and US EPA. Brief of Respondents in Edison Electric Institute, et al v. U.S. EPA, et al. June 8, 2004 @ pg. 45 and pg. 78.

[11] U.S. EPA. Protocol for EPA Approval of New Methods or Alternate Test Procedures for Whole Effluent Toxicity. January, 1999; pg. 1.

[12] 40 CFR 136.4

[13] U.S. EPA. Protocol for EPA Approval of New Methods or Alternate Test Procedures for Whole Effluent Toxicity. January, 1999; pgs. 6 & 7 [emphasis added].

[14] U.S. EPA. National Pollutant Discharge Elimination System Test of Significant Toxicity Technical Document. EPA-833-R-10-004 (June, 2010) pg. 60 [emphasis added].

1.7) Federal regulations require that: *"those who develop or use a modification to an approved (part 136) method must document that the performance of the modified method, in the matrix to which the modified method will be applied, is equivalent to the performance of the approved method. If such a demonstration cannot be made and documented, then the modified method is not an acceptable alternative to the approved method."*[15]

1.8) The Substitute Environmental Document (SED) for the proposed policy acknowledges that: *"for a small number of tests, the TST approach may determine a different outcome than other statistical approaches."*[16] However, the number of times the TST produces a different test outcome is not "small." Data from the State Board's Test Drive study showed that the TST came to a different conclusion in about 8% of all Ceriodaphnia dubia reproduction tests (the single most common endpoint used to evaluate wastewater discharges to freshwater streams in California). In these cases, the TST was nearly twice as likely to label the sample "toxic" compared to the No-Observed-Effect-Concentration (NOEC) metric. Moreover, the Test Drive data indicates that the TST is three times more likely to label the sample as "toxic" compared to the IC-25 procedure that EPA's promulgated method states is the preferred approach for NPDES permitting.[17]

1.9) Federal regulations require all NPDES discharge monitoring reports (DMRs) must be certified by the discharger using the following statement:

> *"I certify under penalty of law that this document and all attachments were prepared under my direction or supervision in accordance with a system designed to assure that qualified personnel properly gather and evaluate the information submitted. Based on my inquiry of the person or persons who manage the system, or those persons responsible for gathering the information, the information is, to the best of my knowledge and belief, true, accurate, and complete. I am aware that there are significant penalties for submitting false information, including the possibility of fine and imprisonment for knowing violations."*[18]

When the results of a toxicity test evaluated using the TST procedure are inconsistent with the conclusion derived when analyzing the same data using EPA's promulgated statistical procedures, then the discharger is unable to "know" which result to "believe." Under such circumstances, permittees cannot be compelled to certify that the effluent is toxic when they don't know which of two valid but contradictory conclusions is "true."[19]

---

[15] 40 CFR 136.6(b)(1)

[16] California State Water Resources Control Board. Draft Staff Report, Including Substitute Environmental Documents, for the Proposed Establishment of the Water Quality Control Plan for Inland Surface Waters, Enclosed Bays, and Estuaries of California; and Toxicity Provisions. October 19, 2018; pg. 127.

[17] U.S. EPA. Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms. EPA-821-R-02-013. Fourth Ed., Oct., 2002; see §9.5.1 on pg. 41.

[18] 40 CFR 122.22(d) [emphasis added].

[19] See, for example, Systech Environmental Corp. v. EPA 555 Fed.3d 1466 (9th Cir., 1995)

Comment #2: The proposed TST procedure cannot be used to evaluate compliance with WET limits in an NPDES permit because it fails to confirm the presence of a valid concentration-response relationship, as required by the promulgated test methods, prior to concluding that a given effluent sample is toxic.

2.1) U.S. EPA has repeatedly and consistently affirmed the essential importance of evaluating the underlying concentration-response relationship when assessing potential toxicity:

> *"The concept of a concentration-response or, more classically a dose-response relationship, is the most fundamental and pervasive one in toxicology…"*[20]

> *"A corollary of the concentration-response concept is that <u>every toxicant</u> should exhibit a concentration-response relationship, given that the appropriate response is measured and given that the concentration range evaluated is appropriate. Use of this concept can be helpful in determining whether an effluent possesses toxicity and in identifying anomalous test results. An evaluation of the concentration-response relationship generated for each sample is an important part of the data review process that should not be overlooked."*[21]

2.2) EPA's promulgated WET test methods <u>require</u> a minimum of five effluent concentrations and a control in order to ensure adequate data to assess the validity of the dose-response relationship.[22]

> *"Effluent chronic toxicity is generally measured using a multi-concentration, or definitive test, consisting of a control and <u>a minimum of five effluent concentrations</u>. The tests are designed to provide dose-response information, expressed as the percent effluent concentration that affects… survival, growth and/or reproduction…"*[23]

> *It is the Agency's policy that <u>all effluent toxicity</u> tests include a minimum of five effluent concentrations and a control."*[24]

[20] U.S. EPA [citing Casarett and Doull, 1975]. Method Guidance and Recommendations for Whole Effluent Toxicity (WET) Testing (40 CFR Part 136). EPA-821-B-00-004. (July, 2000) pg. 4-1

[21] U.S. EPA. Method Guidance and Recommendations for Whole Effluent Toxicity (WET) Testing (40 CFR Part 136). EPA-821-B-00-004. (July, 2000) pg. 4-3 [emphasis added].

[22] U.S. EPA. Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms - 4th Ed. Oct., 2002. EPA-821-R-02-013. See, for example, Item #18 in Table 1 on pg. 76 and Item #17 in Table 3 on pg. 165

[23] U.S. EPA. Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms - 4th Ed. Oct., 2002. EPA-821-R-02-013. §2.2.2 and §2.2.3 @ pg. 5 [emphasis added].

[24] U.S. EPA. Whole Effluent Toxicity: Guidelines Establishing Test Procedures for the Analysis of Pollutants Supplementary Information Document(SID) Oct. 2, 1995 @ pg. 28 [emphasis added].

2.3) The TST procedure is performed by comparing data from only two test concentrations – a control and one effluent concentration. The claim that EPA "neither recommends nor requires review of concentration-response pattern prior to or subsequent to running the TST approach" is misleading.[25] In reality, the TST Technical Document does not discuss dose-response relationships at all. However, this does not imply that such a review is unnecessary because EPA has consistently rejected such an inference:

> *"The agency [EPA] is concerned that single concentration, pass/fail, toxicity tests do not provide sufficient concentration-response information on effluent toxicity to determine compliance. It is the Agency's policy that all effluent toxicity tests include a minimum of five effluent concentrations and a control."*[26]

> *"… the use of pass/fail tests consisting of single effluent concentration (e.g. receiving water concentration or RWC) and a control* ***is not recommended****."*[27]

In fact, U.S. EPA headquarters has explicitly cautioned against the misleading claims being made regarding the need (or lack thereof) for evaluating multiple concentrations when using the TST procedure:

> *"Both the Office of Wastewater Management and Office of Science and Technology have concerns about the memo which mischaracterizes some of the TST document language and endorses a whole effluent toxicity (WET) method approach that is not approved in EPA's promulgated WET test methods (40 CFR Part 136). While the TST document recommends analyzing the data generated from two test concentrations, it still maintains EPA's mandatory test acceptability criteria (TAC) of running WET tests with five concentrations, consistent with EPA's promulgated test methods. A WET test method that uses only two concentrations does not meet the minimum mandatory TAC and therefore requires an alternative test procedure (ATP) before deviating from an EPA test method. It is particularly important to characterize the Headquarters TST document and the Part 136 WET test method requirements in order to appropriately inform California's development of its toxicity policy – including assurance that test data developed under that policy are viewed as valid by complying with EPA's minimum WET test method TACs."*[28]

---

[25] California State Water Resources Control Board. Draft Staff Report, Including Substitute Environmental Documents, for the Proposed Establishment of the Water Quality Control Plan for Inland Surface Waters, Enclosed Bays, and Estuaries of California; and Toxicity Provisions. October 19, 2018; pg. 13.

[26] U.S. EPA. Whole Effluent Toxicity: Guidelines Establishing Test Procedures for the Analysis of Pollutants Supplementary Information Document(SID) Oct. 2, 1995 @ pg. 28. [emphasis added].

[27] U.S. EPA. Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms - 4th Ed. Oct., 2002. EPA-821-R-02-013. See §2.2.2 and §2.2.3 @ pg. 5. See, for example, Item #18 in Table 1 on pg. 76 and Item #17 in Table 3 on pg. 165; [*emphasis in original].*

[28] U.S. EPA. Brennan Ross, Associate Chief of the State and Regional Branch in the Office of Wastewater Management. Email to David W. Smith, Manager of NPDES Permits Office in EPA Region 9 dated 3/18.2013.

2.4) The promulgated method requires evaluation of data from a multiple concentration dilution series to ensure proper interpretation of the results.[29] Without such data, the test results cannot be certified as "true, accurate and complete" on the monthly DMR:

> *"A predictable dose response curve is one of the mandatory requirements for a valid toxicity test. We would never accept analytical results from an instrument producing an abnormal standard curve. The predictable dose response curve, that is increasing toxicity with increasing concentration, is the analogue of the analytical standard curve and is of equal importance in toxicity testing*.[30]
>
> *"The dose response curve is the basis for the validity of a toxicity test. The control serves as the starting point from which the dose response is evaluated. If a dose response is not obtained, then toxicity cannot be inferred."*[31]

2.5) It should be noted that similar comments regarding the invalidity of relying on just two test concentrations, were submitted in 2012 on a previous draft of the state's proposed toxicity policy.[32] State Board staff responded (in 2018) by noting that the current proposal requires dischargers and laboratories to continue running the test method – including the multiple dilution series – as promulgated under 40 CFR Part 136. This, however, is irrelevant because the TST procedure makes no use whatsoever of the test data generated for all but one of the effluent concentrations. So, there is no meaningful difference between the policy proposed in 2018 (which runs multiple effluent concentrations but does not use the data) and the 2012 policy (which ran only a control and just one effluent concentration).

Moreover, as noted earlier, the 2018 Staff Report explicitly claims that one of the benefits of the new approach is that it avoids the complexity and cost of testing or interpreting multiple test concentrations. Forcing dischargers to continue running multiple test concentrations, while deliberately ignoring the data from four of the five effluent exposures in that multi-concentration dilution series, appears to be a transparent attempt to maintain a superficial appearance of consistency with the promulgated method while circumventing federal regulatory requirements prohibiting modifications to test procedures for method-defined analytes. EPA headquarters has already spurned this misguided approach.

2.6) EPA incorporated mandatory review of the concentration-response relationship into the WET method manuals.[33] Ignoring this requirement, by ignoring 60% of the data normally generated during a WET test, constitutes a significant change and an improper modification to the promulgated test methods.

---

[29] U.S. EPA. 67 FR 223, 69962 (Nov. 19, 2002)

[30] U.S. EPA. National Effluent Toxicity Assessment Center, EPA Environmental Research Laboratory in Duluth, MN; NETACommunique, January, 1990 [emphasis added].

[31] Dr. Teresa J. Norberg-King, EPA Environmental Research Laboratory; Permit Review Memorandum to EPA-Region X; June 5, 1989 [emphasis added].

[32] Comment Letter submitted by City of San Bernardino Water Department to Charles Hoppin, Chairman of the State Water Resources Control Board, dated August 21, 2012.

[33] U.S. EPA. Method Guidance and Recommendations for Whole Effluent Toxicity (WET) Testing (40 CFR Part 136). EPA-821-B-00-004. (July, 2000); pg. 4-1 [cited at 67 FR 223, 69971].

Comment #3: The draft policy proposes to supersede a portion of the Santa Ana Regional Water Quality Control Plan where there is no actual conflict between provisions.

3.1) Section III-B-3 of the draft policy states that: "In accordance with Water Code section 13170, except where otherwise noted, the TOXICITY PROVISIONS supersede any Regional Water Quality Control Plans (Basin Plans) for the same waters to the extent of any conflict." Appendix E to the related Staff Report, entitled: Superseded Portions of the Regional Water Board Basin Plans, identifies the specific provisions of each Regional Basin Plan that "conflict" with the proposed policy using a ~~strikeout~~ font. The Staff Report further indicates that the following sentence in the Santa Ana Region's Basin Plan "conflicts" with and will be superseded by the proposed policy:

> *"The Regional Board also encourages the development of scientifically sound toxicity test quality control and standardized interpretation criteria to improve the accuracy and reliability of chronic toxicity demonstrations."*[34]

3.2 There is no explanation provided as to why the above sentence "conflicts" with the proposed policy nor is there any rationale for why this specific provision of the Basin Plan must be superseded. On the contrary, it appears that the entire justification for introducing the new TST procedure rests on the claim that doing so will improve the accuracy and reliability of toxicity test results. If so, it is difficult to comprehend how the Santa Ana region's current basin plan language poses any conflict with the proposed state policy.

3.3 Striking the Basin Plan language that merely seeks to improve the accuracy and reliability of WET testing by encouraging more rigorous QA/QC is inconsistent with the formal certification statement each discharger must make on the monthly DMR. The official certification statement, cited in section 1.9 above, requires that all monitoring data be gathered in accordance with a "system" designed to assure the information is properly evaluated.[35]

> *"The validity or quality of the DMR data is ultimately the permittee's responsibility and is a direct result of the adequacy and functioning of the permittee's self-monitoring program. For the program to function properly, data requirements must be structured so that responses will provide the decision makers with the information necessary to determine compliance and support enforcement."*[36]

Permittees cannot be precluded from developing and applying all reasonable QA/QC measures needed in order to certify that the reported data is "true, accurate and complete" because the Federal Court of Appeals has previously determined that:

---

[34] SWRCB. Division of Water Quality. Draft Staff Report for the Proposed Establishment of the Water Quality Control Plan for Inland Surface Waters, Enclosed Bays and Estuaries of California and Toxicity Provisions. Oct. 19, 2018; pg. 311 (citing Santa Ana Basin Plan, 2008, pg. 6-18).

[35] 40 CFR 122.22(d)

[36] U.S. EPA. NPDES Permit Writer's Guide to Data Quality Objectives; Nov., 1990; p. 3-4

"The possibility of statistical measurement error … deprives the agency of the power to find a violation of the standards, in enforcement proceeding, where the measured departure from them is within the boundaries of the probable measurement error."[37]

3.4 Striking the subject language from the Santa Ana Region's Basin Plan is inconsistent with EPA guidance which explicitly endorsees the application of additional QA/QC measures to minimize analytical variability, especially in WET testing:

"Four main components of WET tests afford opportunities to control and minimize variability within tests and within and between laboratories: 1) quality control [QC] procedures; 2) experimental design; 3) test power; and 4) test acceptability criteria [TAC] beyond the minimum requirements specified in EPA's WET test methods."[38]

"State and Regional permit writers should develop their own Data Quality Objectives (DQOs) that are tailored to requirements of their permitting policies. At a minimum, those DQOs should meet and be consistent with the Federal DQO guidelines. The issues discussed in this document provide NPDES permit writers with a basis for developing DQOs and a sensitivity toward the DQO factors: precision, accuracy, comparability, representativeness and completeness."[39]

"Regulatory authorities should ensure that statistical procedures and test methods have been properly applied to produce WET test results. Evaluating other factors and data, such as biological and statistical quality assurance, and ensuring that test conditions and test acceptability criteria (TAC) have been met would be prudent. Regulatory authorities should develop a QC checklist to assist in evaluating and interpreting toxicity test results."[40]

EPA has published two guidance manuals, both of which were cited in the promulgated WET test methods, describing numerous additional QA/QC tools that can be employed when interpreting toxicity data. These manuals also provide several examples from WET implementation programs in other states to support such an approach.[41] Therefore, the Santa Ana Basin Plan provision in question does NOT conflict with the proposed policy and the Staff Report should be revised to eliminate any such implication.

[37] Amoco Oil Co. v. EPA, 501 F.2d 722 (D.C. Cir. 1974)
[38] U.S. EPA. Understanding and Accounting for Method Variability in Whole Effluent Toxicity Applications Under the NPDES Program (EPA 833-R-00-003); June, 2000; p. 5-2 [emphasis added].
[39] U.S. EPA. NPDES Permit Writer's Guide to Data Quality Objectives; Nov., 1990; p. 1-1
[40] U.S. EPA. Understanding and Accounting for Method Variability in Whole Effluent Toxicity Applications Under the NPDES Program (EPA 833-R-00-003); June, 2000; p. 7-4
[41] 67 FR 223, 69971 (Nov. 19, 2002)

Comment #4: The proposed TST approach does not significantly improve the accuracy of tests that rely on Ceriodaphnia dubia reproduction as a primary measure of potential effluent toxicity nor does it materially reduce uncertainty about the reported results.

4.1 The draft Staff Report states that:

> *"The TST approach improves upon the traditional hypothesis tests used to assess aquatic toxicity by establishing Regulatory Management Decisions (RMDs) and through the reversal of the null and alternative hypothesis. The TST approach's RMDs provide an unambiguous measurement of a test concentration's toxicity, while low false positive and false negative rates prove more statistical power to correctly identify a test concentration as toxic or non-toxic."*[42]

This conclusion appears to be largely based on results reported from the State Board's "TST Test Drive" study.[43] However, the pie-chart summaries presented in Figures 5-1 and 5-2 of the draft Staff Report improperly combine all of the data from different test methods (i.e. acute and chronic), and different receiving water regimes (i.e. marine and freshwater), and numerous different species, and different biological endpoints (lethal and sub-lethal). The fact that the TST approach may work well when evaluating survival rates of a marine species such as Mysid shrimp is completely irrelevant to how the TST approach performs when used to evaluate toxicity based on a sub-lethal endpoint in a freshwater species (e.g. Ceriodaphnia dubia reproduction). Data from the different species and endpoints cannot be combined to support the conclusion that, overall or on average, the TST method is "better." Such an approach merely obscures the true performance of the TST approach for each test method where it will be applied in the context of NPDES compliance monitoring: one species/endpoint at a time.

4.2 The Test Drive Study did not compare TST performance to the point estimate techniques that the promulgated methods identify as "preferred" for the purpose of evaluating effluent toxicity in the context of the NPDES permitting program.[44] By design, the IC-25 endpoint will identify all tests where the measured adverse effect is greater than 25% as toxic and all tests where the measured adverse effect is less than 25% as non-toxic. These results are superior to both hypothesis-testing techniques (NOEC and TST).

[42] California State Water Resources Control Board. Draft Staff Report, Including Substitute Environmental Documents, for the Proposed Establishment of the Water Quality Control Plan for Inland Surface Waters, Enclosed Bays, and Estuaries of California; and Toxicity Provisions. October 19, 2018; pg. ix.

[43] California State Water Resources Control Board. The Effluent, Stormwater and Ambient Toxicity Test Drive Analysis of the Test of Significant Toxicity (TST). 2011 [cited and summarized on pgs. 57-60 of the draft Staff Report].

[44] U.S. EPA. Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms. EPA-821-R-02-013 - Fourth Ed., October, 2002; §9.5.1 @ pg. 41.

4.3 Data from the Ceriodaphnia dubia reproduction tests analyzed in conjunction with the Test Drive Study clearly show that the TST procedure does not produce "unambiguous" results:

(a) There were1,095 Ceriodaphnia dubia reproduction tests reported in the TST Test Drive database recently made available on the State Board's website.[45]

(b) 19 of the 1,095 tests failed to meet EPA's mandatory Test Acceptability Criteria for minimum control performance and were excluded from further analysis. That leaves 1,076 valid Ceriodaphnia dubia reproduction tests.

(c) The NOEC and TST appear to disagree on whether the sample was toxic in 88 (8%) of the 1,076 valid Ceriodaphnia dubia reproduction tests.

(c) In 31 of these 88 discordant cases the TST passed and the NOEC failed. However, in 15 of these 31 tests there was less than a 13% difference between the effluent-exposed group and the control group. EPA's method manual states that such small differences should not be counted as actual test failures even if there is a statistically-significant difference (see §10.2.8.2.5 @ pg. 51 of EPA's 2002 Chronic Freshwater Method Manual). After making this adjustment only there were only 16 tests where the NOEC failed and the TST passed. Note: it appears that the Test Drive study failed to make this adjustment and, contrary to the promulgated method, improperly counted these marginal results as actual toxicity.

(d) There were 57 tests where the TST failed and the NOEC passed; 5 of these 57 tests had less than a 10% difference between the effluent-exposed group and the control group. EPA's TST guidance suggests that such small differences may be anomalous and can be considered potential Type-I errors (i.e. failure to properly reject the null hypothesis). Therefore, to be fair, these five tests were also excluded from further consideration. This leaves 52 tests where the TST failed and the NOEC passed.

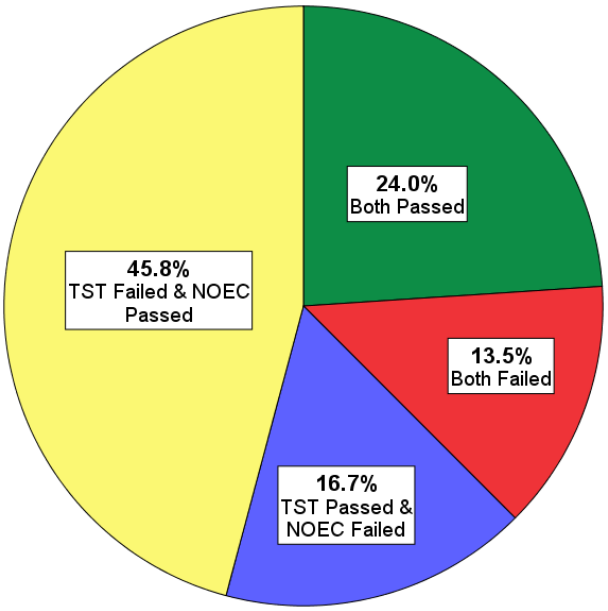
(e) After removing the marginal NOEC and TST failures, in accordance with EPA's recommendations, there are a total of 68 tests where the NOEC and TST methods disagreed on whether a given sample was toxic or not. However, the Percent Effect exceeded the Regulatory Management Decision threshold (i.e. 25% difference) in only 8 (12%) of these 68 tests. This indicates that, at worst, the NOEC metric failed to accurately identify the presence of toxicity in less than 1% of the 1,076 valid Ceriodaphnia dubia reproduction tests evaluated during the "Test Drive" study. Such results do not suggest that the promulgated NOEC method is seriously underestimating the true incidence of potential effluent toxicity. EPA's preferred point estimation procedure would have identified all 8 of these samples as toxic.

[45] There were no flags or fields in the downloaded spreadsheet to indicate which of the 1,095 Ceriodaphnia dubia reproduction tests the state used and which were excluded. There also appear to be numerous data entry errors and other miscalculations in the worksheet. These errors and omissions should be corrected and the worksheet reposted with a new 30-day review and comment period. Until then, we have no choice to analyze the TST Test Drive data just as it was when we downloaded it from the State Board's website.

(f) In the 60 tests where the NOEC and TST disagreed on whether the sample was toxic or not, and the percent effect was LESS than the Regulatory Management Decision threshold, the TST was nearly three times more likely to label such samples as "toxic" compared to the NOEC (44 vs. 16, respectively). In all 60 of these tests, average reproduction for the effluent-exposed organisms was MORE than 75% of that reported for the control group. However, the TST procedure will require that dischargers report and certify such samples as "toxic" despite the fact that, statistically, there is less than a 50% chance that these samples were actually toxic. EPA's preferred point estimation procedure would have concluded that all 60 of these samples were "Not Toxic."

(g) Since EPA's preferred point estimation technique identifies all tests that exhibit more than a 25% effect as toxic, and the promulgated method states that Ceriodaphnia dubia tests showing less than a 13% reduction in average reproduction should not be deemed "toxic," the critical question for the State Board is: how does the TST technique compare to the NOEC procedure for those tests that exhibit effects larger than 13% but less than 25%. In the Test Drive Study, only 96 of the 1,076 valid tests fell into this "zone of uncertainty." The pie chart below summarizes the results for these 96 tests (see Fig. 1):

Fig. 1: Discordant Results for C. dubia Reproduction Tests in the Test Drive Study



In these tests, where intrinsic natural biological variability and the resulting uncertainty matter most, the NOEC and IC25 disagreed as to whether the sample was toxic or not more than 62% of the time. And, in such cases, the TST was nearly three times more likely to label the sample as toxic compared to the NOEC. For the Ceriodaphnia dubia reproduction endpoint, the TST does not produce "unambiguous" conclusions - especially where the actual Percent Effect observed in a given test is less than 25%.

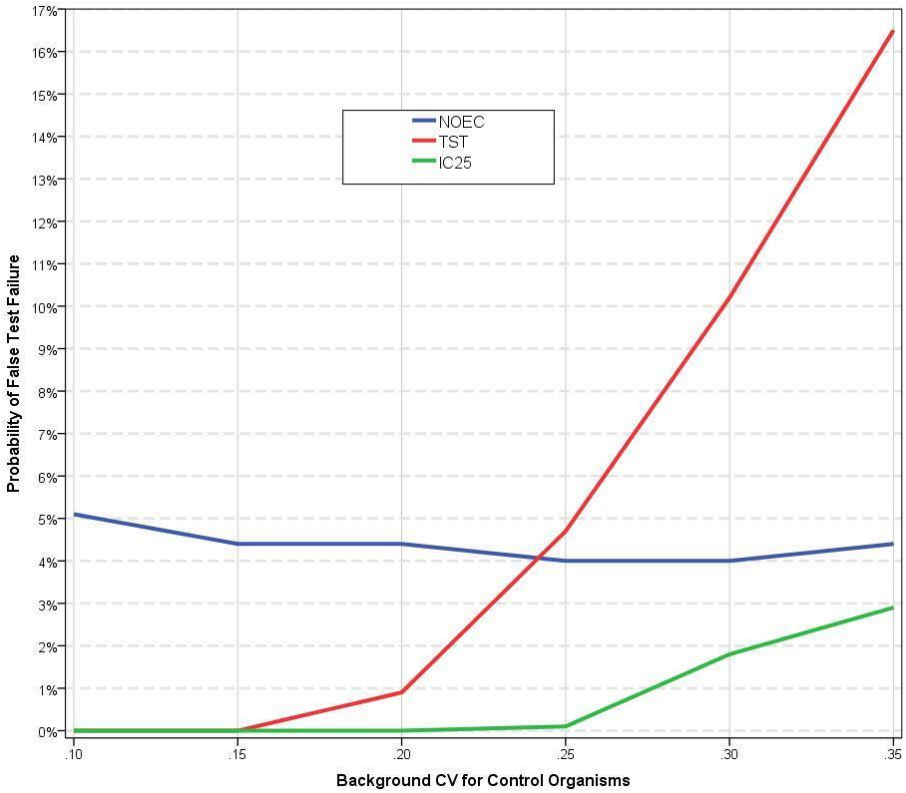
4.4 Based on the results from a computer simulation study described in written comments previously submitted to the State Board, the accuracy of the TST method degrades rapidly when the inter-replicate coefficient-of-variation (CV) is greater 20%.[46] The table and chart below provide a visual summary of the simulation study data that was previously presented in Risk Sciences' 2012 comment letter.

Effect of CV on False Failure Rates for C. dubia Reproduction

CV	NOEC	TST	IC25
10%	5.1%	0%	0%
15%	4.5%	0%	0%
20%	4.5%	0.9%	0%
25%	4.1%	5.3%	0.3%
30%	3.8%	10.9%	1.8%
35%	4.3%	17.4%	3.1%

(N of replicates = 10)

Fig. 2: Probability of False Failure as a Function of Control Variability for C. dubia Reproduction



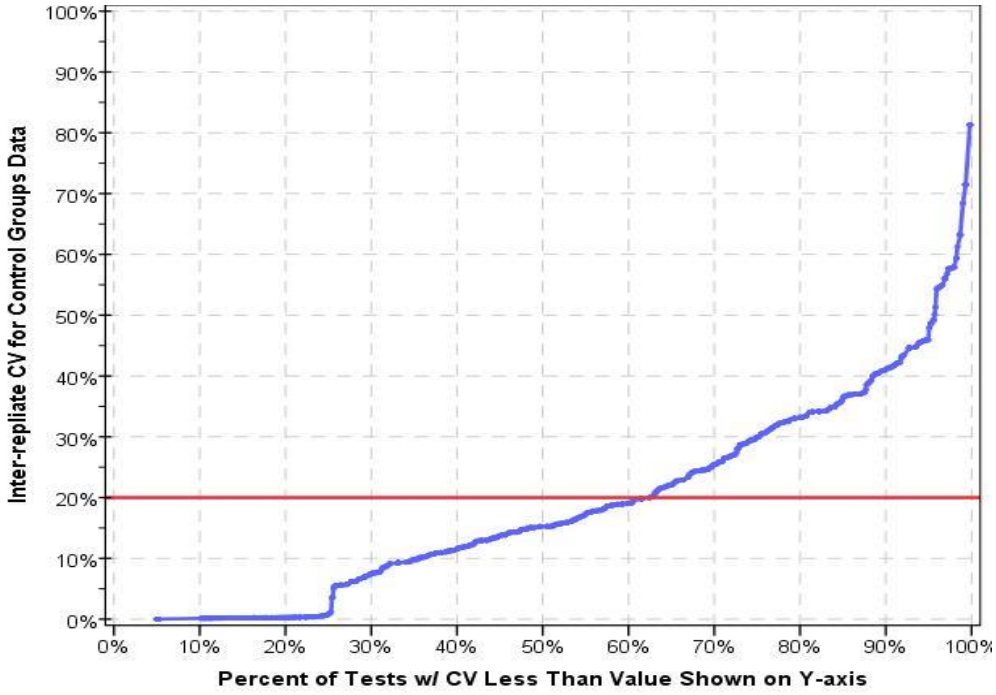
[46] Risk Sciences. Comment Letter to Charles Hoppin, Chairman of the State Water Resources Control Board, re: Draft Policy for Toxicity Assessment and Control. August 21, 2012.

In the Response-to-Comments document, State Board staff dismissed Risk Sciences' analysis by claiming it was based on an outdated EPA study and obsolete assumptions about test variability.[47] This is not true; an entirely new study was submitted showing how the TST performed over a wide range of variability. The response given simply fails to address the comment submitted or the actual evidence offered to support that comment.[48]

Staff's response also asserts that laboratory performance had improved substantially over the years. No evidence was offered to support this claim and the TST Study shows it is untrue. In several cases where the TST "failed," despite the fact that the average reproduction for effluent exposed organisms was actually greater than 75% of that in the control group, the TST Study report explicitly blamed excessive data variability.[49] This demonstrates , even if laboratory performance has improved over the years, intrinsic natural biological variability still remains a significant problem when applying the TST procedure to real-world test data.

Review of control performance data from the Ceriodaphnia dubia reproduction tests analyzed during the TST Test Drive Study shows that the inter-replicate coefficient-of-variation (CV) exceeded 25% about one-third of the time. In one-fourth of the tests the CV was greater than 30% and 10% of the time the CV was greater than 40% (see Fig. 3, below).

Fig. 3: CV for C. dubia Reproduction in Control Group Data from the TST Test Drive Study



[47] State Water Resources Control Board. Response to Comments on the 2012 Draft Policy for Toxicity Assessment and Control. October 26, 2018 (see Comment 36.2 on pg. 72).

[48] For the reasons outlined above, Risk Sciences 2012 comment letter is being resubmitted (by reference) as a new comment on the draft policy now before the State Board for consideration.

[49] California State Water Resources Control Board. The Effluent, Stormwater and Ambient Toxicity Test Drive Analysis of the Test of Significant Toxicity (TST); 2011; pg. 12.

It is when the CV is greater than 25% that the TST procedure is most likely to lead to false conclusions regarding potential toxicity in the effluent sample. It should be noted that the computer simulation studies, performed by TetraTech in order to develop the TST procedure, used synthetic datasets with an assumed Coefficient of Variation of less than 20%. Data from the Test Drive Study reveals that the real-world CV is greater than the assumed value nearly 40% of the time (see Fig. 3).

The Staff Report suggests that dischargers can avoid such problems by simply increasing the number of replicate organisms used in each WET test.[50] The Test Drive study provides examples of how doing so would change some TST results from "fail" to "pass."[51] This demonstrates that the proposed TST procedure does not produce "unambiguous" results when used with the promulgated WET test methods. These methods must be modified to increase the number of replicates in order to ensure that any compliance determinations made using the TST are, in fact, correct.

4.5 Uncertainties associated with using the TST procedure make it virtually impossible for permittees to certify some WET test results on a Discharge Monitoring Report (DMR).

According to data from the Test Drive Study, the TST technique and the promulgated NOEC method support inconsistent conclusions about the presence or absence of toxicity in nearly 9% of all Ceriodaphnia dubia reproduction tests that were evaluated. In those cases where average reproduction among the effluent-exposed organisms was greater than 75% of the control performance, there is serious question as to why the TST failed, especially where EPA's preferred point estimation method identifies such samples as "non-toxic." When the IC25 passes and the TST fails, there is less than a 50% chance that the effluent sample actually violated the RMD threshold. However, the sample continues to be "presumed toxic" and the permittee is required to certify that the TST a test "failed" despite the fact it is more likely than not that the true effect is less than 25% effect (i.e. not toxic). A discharger cannot be compelled to certify, to the best of their "knowledge and belief," that the effluent is toxic and the TST result is "true" when the promulgated statistical technique "preferred" by EPA supports exactly the opposite conclusion. Mandating such a certification, despite all valid evidence to the contrary, is both unreasonable and illegal.[52]

> *"Use of the TST approach does not preclude the use of EPA's Technical Support Document (TSD) approach for analyzing valid WET data or another scientifically defensible approach that is sufficient to meet the statutory and regulatory requirements."*[53]

[50] California State Water Resources Control Board. Draft Staff Report, Including Substitute Environmental Documents, for the Proposed Establishment of the Water Quality Control Plan for Inland Surface Waters, Enclosed Bays, and Estuaries of California; and Toxicity Provisions. October 19, 2018; pgs. 56-58.

[51] California State Water Resources Control Board. The Effluent, Stormwater and Ambient Toxicity Test Drive Analysis of the Test of Significant Toxicity (TST). 2011; Table 3-6 @ pg. 13.

[52] See, for example, Systech Environmental Corp. v. EPA 555 Fed.3d 1466 (9th Cir., 1995); see also Amoco Oil Co. v. EPA, 501 F.2d 722 (D.C. Cir. 1974); see also 33 U.S.C. 1319(4)

[53] U.S. EPA. NPDES Test of Significant Toxicity Implementation Document. EPA-833-R010-003; June, 2010; pg. 1