Western States Petroleum Association
Credible Solutions • Responsive Service • Since 1907

Kevin Buchan
Senior Coordinator, Bay Area and State Water Issues

VIA ELECTRONIC MAIL

January 21, 2011

Chair Hoppin, and Members of the Board
State Water Resources Control Board
1001 I Street
Sacramento, CA 95814

Subject: Policy for Toxicity Assessment and Control

Dear Chair Hoppin, Members of the Board

The Western States Petroleum Association (WSPA) is a non-profit trade association representing twenty-six companies that explore for, produce, refine, transport and market petroleum, petroleum products, natural gas and other energy supplies in California, Arizona, Nevada, Oregon, Washington and Hawaii. WSPA appreciates the opportunity to comment upon the State Board's proposed draft Policy for Toxicity Assessment and Control.

WSPA recognizes and appreciates the effort that State Board staff has put into the development of the proposed Draft Policy for Toxicity Assessment and Control (the Draft Policy). However, WSPA has a number of serious concerns regarding the Draft Policy. We summarize our concerns in this cover letter, and provide greater detail in the attachment following it.

1. **Implementation as numeric effluent limitations is inappropriate and problematic.** The Draft Policy establishes numeric objectives for chronic and acute toxicity; these objectives are intended for use as numeric effluent limitations in NPDES permits. Exceedance of numeric effluent limitations derived from the Draft Policy would constitute a violation of the permit.

   Unlike chemical analyses, toxicity tests measure responses of certain test organisms, and can be influenced by numerous factors other than and in addition to effluent toxicity. For these reasons, failure of any single toxicity test should not be automatically

considered to be a violation but rather should trigger further investigation to determine if the effluent is toxic and/or to identify a toxicant(s).

2. **The use of USEPA's TST method, and the application of toxicity requirements as numeric effluent limitations is inappropriate.** The Draft Policy requires the use of the TST approach to test for whole effluent toxicity (WET), even though the TST method was established by USEPA as guidance and has not been through the public review and comment process. The statistical measures and hypotheses of the Draft Policy (and of the TST) assume that an effluent is toxic unless testing is able to demonstrate that the effluent is in fact not toxic—a reversal of the "presumption of innocence," and a significant departure from traditional practice.

   To evaluate the TST method, we applied it to USEPA WET blank data, which by definition are non-toxic. Our analyses showed that the TST method falsely indicates toxicity in these samples at a rate of 15 % for chronic toxicity tests using *Ceriodaphnia dubia* (*C. dubia*) for reproduction.

   We also applied the TST method to other datasets for *C. dubia* for reproduction and found a significantly higher rate of toxicity than with methods currently in use. These toxicity findings are most likely due to inherent variability in sublethal chronic toxicity test methods, to inter-laboratory differences, to variability caused by sample matrices (e.g., hardness, pH, TDS of receiving water), and variability inherent in the hypothesis testing in the TST method —***not*** to actual toxicity in the samples themselves. For these reasons, use of the TST method will provide no improvement in terms of protecting the aquatic environment as compared to currently existing methods (e.g., NOEC, IC25).

   Our analyses indicate that the use of the TST method will lead to an increase in false violations (i.e., determining that an effluent sample is toxic when the effluent is, in truth, not toxic). This, in turn, will lead to the unnecessary expenditure of significant resources by the State and Regional Water Board and the discharger community to respond to non-toxic, false indications of toxicity, in addition to potentially unwarranted enforcement actions and unjustified 303(d) listings.

   Similar analyses should be performed for other species and other test endpoints, which should be used by the State only if significantly lower rates of false toxicity are found to exist. In any case, the TST method should not be used to derive numeric effluent limitations. With lower rates of false toxicity, the TST method could potentially be used as triggers for additional testing and investigation in conjunction with a narrative objective for toxicity.

3. **The reasonable potential analysis (RPA) of the Draft Policy will result in unnecessary application of effluent limitations.** The Draft Policy results in a finding of reasonable potential (i.e., the determination that a discharge has the potential to cause or contribute to an excursion above a water quality standard, and thus requires an effluent limitation) under either of two conditions: (1) if an effluent sample fails the TST method or (2) if the percent effect (i.e., the difference between responses of the effluent sample and the control) is greater than 10%. Because of the variability inherent in toxicity testing, particularly for sublethal, chronic toxicity endpoints, the second condition would be expected to be frequently exceeded. Our analyses demonstrated a false failure rate of the RPA at 25% for chronic toxicity tests using *C. dubia* for reproduction

from USEPA WET blank data. The State Board should not consider for adoption a proposed method with such a high false failure rate.

4. **The Draft Policy unnecessarily increases reliance on animal testing.** If monthly routine monitoring is required by an NPDES permit for the most sensitive species, a minimum of 96 fathead minnows (or 240 *C. dubia*) will be used in annual testing. If accelerated monitoring is triggered by false violations at the assumed rate of 5% for fathead minnows (or at the rate of 15% for *C. dubia* as found in the EPA WET blank reproduction data), then 144 fathead minnows (or 1,080 *C. dubia*) will be used annually in toxicity testing. This equates to 720 fathead minnows (or 5,400 *C. dubia*) during a 5-year permit term. WSPA members have adopted global directives to reduce the number of animals used in testing, and these global commitments would be undermined by this Draft Policy if adopted.

5. **The proposed Draft Policy should not be considered until the TST Method has been peer reviewed and adopted through a formal rule-making process.** The State Board's definition of whole effluent toxicity using the entirely new TST method constitutes a change in water quality standards. The proposed Draft Policy and its contents must be adopted through a formal rule-making process, and the State Board must comply with the requirements in California Water Code Sections 13241 and 13242. The TST method, upon which the Draft Policy is based, has not been adopted through a public process.

6. **The TST method is not approved for use under current legal requirements.** 40 CFR Part 136 contains guidelines establishing test procedures for the analysis of pollutants. The TST method is new and has not been approved as meeting these guidelines. No federal register notices were released concerning the TST guidance. No peer review that meets California requirements was conducted by USEPA. No estimates were made of inter-laboratory test precision, which is required for any new method. It has not yet been authorized as an alternate test procedure pursuant to 40CFR 136.5 Approval of Alternate Test Procedures.
7. **The cost analysis significantly underestimates the costs and environmental impacts of the Draft Policy.** The economic analyses contained in the Staff Report for the Draft Policy underestimate the likely monitoring costs. Neither the economic and environmental impact analyses considered the reasonably foreseeable costs of compliance. In order to comply with the provisions in the Draft Policy, we are concerned that additional treatment facilities may be required (potentially including nitrification, disinfection by UV/ozone, activated carbon, and/or reverse osmosis), even if the findings of toxicity are false.

8. **Expansion of whole effluent toxicity testing to stormwater discharges is invalid.** The proposed Policy expands whole effluent toxicity testing to stormwater dischargers even though this expansion is unsupported by appropriate studies or data collection. This expansion would be expected to result in a significant increase in enforcement actions and related appeals.

These and other concerns are discussed in greater detail in the attachment to this letter.

In summary, WSPA strongly recommends that the State Board not adopt the numeric objectives or use of the TST method in the Draft Policy to derive numeric effluent limitations.

Existing methods and data support the continued use of narrative objectives with accelerated monitoring and toxicity reduction evaluation (TRE) triggers to address effluent toxicity. These methods have been effectively implemented in California for several years, are consistent with guidance from USEPA, and are supported by recognized national and regional experts. It may be appropriate to use the TST methods as one component of a Policy based on narrative objectives, but only after significant additional analysis and only for those species and endpoints that have reasonable rates of finding false toxicity.

WSPA looks forward to working with the State Water Board and its staff on future revisions to the Draft Policy. Please contact me at your convenience if you have any questions regarding our comments.

Sincerely,

Attachment: Detailed Comments

1415 "L" Street, Suite 600, Sacramento, California 95814
(916) 498-7755 ● FAX: (916) 444-5745 ● Kevin@wspa.org ● www.wspa.org

ATTACHMENT
Detailed Comments provided by the Western States Petroleum Association (WSPA)
January 21, 2011

Introduction

The State Water Resources Control Board's Draft Policy for Toxicity Assessment and Control (Draft Policy) is based upon the United States Environmental Protection Agency's (USEPA) recently released Test of Significant Toxicity (TST)[1]. The Draft Policy would establish numeric objectives for whole effluent toxicity (WET), which the Staff Report for the Draft Policy states are "simply a concise statement of several elements in USEPA's TST document."[2] The Draft Policy requires use of chronic sub-lethal toxicity tests and the TST method in order to determine whether an effluent sample exceeds the numeric objectives established by the Draft Policy. Under the Draft Policy, exceedance of the numeric objectives placed into NPDES permits pursuant to this policy would be considered a violation of the permit, and would trigger further processes such as accelerated monitoring and a Toxic Reduction Evaluation (TRE). The Draft Policy also includes provisions for determining reasonable potential; under the Draft Policy, an effluent sample is found to have "reasonable potential" to cause or contribute to an exceedance of chronic toxicity objectives either if the sample fails the TST method or if the percent effect is greater than 10%, even if the sample passes the TST method test.

1. Numeric objectives are inappropriate and unnecessary for toxicity assessment.

The proposed Draft Policy establishes numeric objectives for whole effluent toxicity. The failure to meet the numeric objectives, once implemented in an NPDES permit as numeric effluent limitations, would be considered a permit violation. Unlike chemical analyses, toxicity tests measure responses of certain test organisms and can be influenced by numerous factors other than and in addition to effluent toxicity. For these reasons, failure of any single toxicity test should not be automatically considered to be a violation but rather should trigger further investigation to determine if the effluent is toxic and/or to identify a toxicant(s). Additional detail is provided below.

Toxicity tests are inherently variable, and the natural biological variability of test organisms undermines the reliability and reproducibility of toxicity test results. Toxicity tests use certain living organisms, as specified in standard toxicity test methods, and measure the responses of these organisms. These organisms can be impacted by numerous factors other than the actual toxicity of effluent, including temperature, light, seasonality, trace elements, total dissolved solids (TDS), and the ionic composition of a water sample. As one example, a large amount of variability in test results can occur as a result of differences between the natural ionic chemistry of an effluent sample and of control waters. Variations in ionic strength and ionic composition may interfere with the conduct and interpretation of toxicity test results[3]. A study published by the Water Environment Research Foundation (WERF) found that more

[1] USEPA, 2010. National Pollutant Discharge Elimination System Test of Significant Toxicity Implementation Document. June.
[2] State Water Resources Control Board, 2010. Staff Report for "Draft Policy for Toxicity Assessment and Control" at p. 65.
[3] Several examples serve to illustrate the problems that may be encountered as a result of differences in ionic strength and ionic composition:

- Example-A: In some cities, local water supplies are very soft. Extreme low hardness is known to adversely impact the normal rate of reproduction in *Ceriodaphnia dubia (C. dubia)*. In this instance, it is the absence of essential

than half of all WET test variability was due to factors other than actual effluent quality (such as choice of laboratories, the particular laboratory technician, the health of the test cultures, and even the season of the year).[4]

In a guidance documents, USEPA warns that:

> *"The interpretation of the results of the analysis of the data from any of the toxicity tests described in this manual can become problematic because of the inherent variability and sometimes unavoidable anomalies in biological data."*[5] (emphasis added)

By itself, analytical variability does not invalidate the toxicity test methods. However, federal courts have ruled that USEPA cannot rely on results that fall within the expected error band of the test to find that permit violation has occurred:

> *"[Analytical variability]…deprives the agency of the power to find a violation of the standards, in enforcement proceedings, where the measured departure from them is within the boundaries of the probable measurement error."*[6]

The U.S. Court of Appeals has ruled that such variability must be accounted for during the permitting process:

> *"[Petitioner's] concern is that some discharge permits may specify an acceptable non-zero level of toxicity, which the effluent may not exceed, and that the WET tests have the potential to produce arbitrary permit violations… This is certainly a problem for which EPA's system must account."*[7]

There are no means by which the discharger can eliminate errors or variability that is unrelated to actual effluent quality. The false violations are a normal and unavoidable limitation of the effluent toxicity testing. This is why the proposed numeric objectives, which specify that a single toxicity test failure constitutes a permit violation, are not appropriate for tests that evaluate effluent toxicity. Any Policy developed for toxicity must recognize and address this mathematical fact. As such, the State Board should not necessarily assume all test failures are prima facie evidence of toxicity or constitute a permit violation. This is particularly important where dischargers are required to perform a very large number of

---

elements rather than the presence of harmful pollutants that may lead to mistaken conclusions about effluent "toxicity."

- Example-B: Storm water samples routinely fail chronic toxicity tests because rainfall is naturally low in conductivity and pH. Generally speaking, rain water is ill-suited for optimum growth of the standard species used as indicator organisms in toxicity testing.
- Example-C: Many states in the arid southwest rely extensively on groundwater to meet public demand for potable water supplies. The natural chemistry of groundwater is distinctly different from surface freshwater. The balance of ions in groundwater may interfere with natural growth and reproduction in the standard test species.

[4] Water Environment Research Foundation, *Whole Effluent Toxicity Testing Methods: Accounting for Variance*, Report #D93002 (1999) at pp. 3-3 & 3-4.

[5] USEPA. Short-Term Methods for Estimating the Chronic Toxicity of Effluent and Receiving Water to Freshwater Organisms, Fourth Ed. EPA-821-R-02-013. October, 2002. Section 9.4.1.1 at p. 39.

[6] *Amoco Oil Co. v. Environmental Protection Agency*, 501 F.2d 722 (D.C. Cir. 1974)

[7] Edison Electric Institute, et al v. Environmental Protection Agency; Case No. 96-1062; Dec. 10, 2004 at p. 8.

tests, as could occur in a monthly monitoring regime using three different species to determine if reasonable potential is exhibited.

USEPA has attempted to establish the TST method such that the error rate (i.e., the rate of finding false violations, also known as the rate of false positives) is set at 5%.[8] (As noted below, our analysis indicates that the rate of finding toxicity in blank samples, which are by definition non-toxic, is actually much higher.) If we assume for the moment that USEPA has correctly specified an error rate for false violations of exactly 5%, it can easily be shown that it is tremendously difficult to demonstrate compliance with a permit requiring monthly sampling. While the risk that any individual toxicity test may falsely determine that a non-toxic effluent sample is toxic appears relatively low (e.g. 5%), the probability of incorrectly finding toxicity is quite high when dischargers perform a large number of tests (e.g., over the life of a permit). If a discharger is required to perform monthly chronic testing for one species, the discharger will have to perform at least 60 separate tests during any given 5-year permit period. The probability of passing all 60 tests is less than 5% (see Table 1), even if the effluent is chemically identical to the non-toxic control water used by the laboratory. Table 1 illustrates the probability of false violation error for this sampling scenario. It is a common misunderstanding that test error rates never exceed a set value by USEPA and are thus always very low. However, test error rates vary as a function of the number of tests performed, even though the average error rate remains fixed at the set value[9].

The Draft Policy would require accelerated testing whenever results from a single effluent sample are found to be toxic. Thus, it is likely that a discharger would be required to conduct more than the 60 tests at a minimum for a permit that requires monthly testing for a single species. Subsequently, the actual number of false violations would likely increase as the discharger performs additional accelerated tests as required by the Draft Policy after the first test failure occurs. As described above, the number of inaccurate test failures would also be much higher if ambient hardness or TDS concentrations cause significant interference with the method; this is particularly important where the natural ionic chemistry of an effluent sample may interfere with the conduct and interpretation of toxicity test results.

Factors that undermine the reliability of the toxicity testing are discussed further in the remainder of this section.

[8] USEPA, 2010. National Pollutant Discharge Elimination System Test of Significant Toxicity Implementation Document at p. vi.

[9] The probability of observing an exact number of errors (X) in a number of analyses (N) when the risk of error for any single test is P is: $\{N!/[(N-X)!*X!]\}*[P^X*(1 - P)^{N-X}]$; Rice, J. A. (2007). Mathematical statistics and data analysis. Belmont, CA, Duxbury press.

Table 1. Probability of the number of false violation errors with the assumed 5% error rate.

(As shown below, note that the actual rate at which false toxicity occurs is significantly higher than 5%).

No. of false violation errors occurred in 60 statistical analyses	Probability of observing **exactly** as many false violation errors	Probability of observing **at least** as many false violation errors
0 in 60	4.6%	-
1 in 60	14.5%	95.39%
2 in 60	22.6%	80.84%
3 in 60	23.0%	58.26%
4 in 60	17.2%	35.27%
5 in 60	10.2%	18.03%
6 in 60	4.9%	7.87%
7 in 60	2.0%	2.97%
8 in 60	0.7%	0.98%
9 in 60	0.2%	0.28%

2. The TST method is scientifically unproven and provides no additional protection for the aquatic environment as compared to existing toxicity test methods.

A. The TST method inappropriately reverses the presumption of innocence

The Draft Policy reverses the presumption of innocence. The statistical hypothesis used in the TST and in the Draft Policy "flips" the standard assumption that an effluent is non-toxic unless shown otherwise. The Draft Policy specifies that a discharger must assume its effluent is toxic and requires the use of the TST method to prove that it is not. Previously the discharger was given the benefit of the doubt in the face of statistical uncertainty. Now, such uncertainty will work against the discharger since it is extremely difficult to prove the negative.

B. False violation error rates increase using the TST method

The mathematics of the TST method significantly increase the rate at which non-toxic effluent samples will be falsely declared to be toxic, contrary to the claims of USEPA and State Board staff that the TST method would significantly reduce statistical uncertainty and the adverse effects of variability.

For example, blank data from USEPA's own toxicity test data for *Ceriodaphnia dubia* (*C. dubia*) show that 15% (i.e., four of 27 samples; see Table 2) of blank samples, which are by definition non-toxic, will incorrectly be deemed toxic using the TST method.

As shown in Table 2, only one out of 27 samples is falsely determined to be toxic by the current 40 CFR 136 methods [i.e., no observable effect concentration (NOEC) and 25 percent inhibition concentration (IC25)], while four of 27 samples are falsely determined to be toxic by the TST method.

Further, Table 2 shows that an additional three blank samples passed the TST test but exhibited greater than a 10% difference in response compared to the control samples. Thus, under the Draft Policy, seven of 27 samples, or 26% of the USEPA blank samples, would have failed the RPA test.

A false violation error rate of 15% means that a discharger, who is required to conduct a monthly test using one species (60 tests for a permit period of five years), will face a 95% probability of being determined to be in false violation of their permit at least once every year during the 5-year permit period (5 in 60 tests; see Table 3), and a 99.99% probability of being determined in violation at least once during the 5-year permit term (1 in 60 tests; see Table 3). Dischargers would incur significant liability under the Clean Water Act and state law for these false violations.

WSPA is concerned that the excessive numbers of false test failures would overwhelm State and Regional Boards and result in unnecessary and unjustified enforcement actions and citizen lawsuits. Enforcement resources would be diverted away from real water quality violations, and dischargers would be put in the untenable position of resolving problems that did not exist.

False determinations of toxicity impairment would also lead to unnecessary and unjustified 303(d) listings. The State's 303(d) Listing Policy[10] at Table 3-1 details the requirements for listing a water on the 303(d) list for toxicants. The requirements are based on the overall sample size and the number of samples that exceed water quality objectives. The table provides that a waterbody must be listed if 2 or more samples exceed objectives for a sample size of 24 or fewer samples. In essence, the table requires listing if 8% to 12% of all available samples exceed water quality objectives. The TST method in the Draft Policy has a 15% rate of finding false toxicity. This means that the TST approach will result in the listing of the majority of water bodies in the State, even if they are actually non-toxic. The probability of falsely listing a waterbody for toxicity impairment (i.e., of observing at least two false TST exceedances in 24 samples) is 89% at the error rate of 15%.

The analysis presented in Table 2 was conducted using USEPA blank data for *C. dubia* for the reproduction endpoint. Based on this analysis, it appears that the reproduction endpoint for *C. dubia* leads to an unacceptably high rate of false toxicity. We have performed preliminary analyses for other endpoints (e.g., fathead minnow) using USEPA blank data; several of these analyses also show elevated rates of "false toxicity" compared to the IC25 and NOEC methods. Clearly, a robust analysis of the performance of various species and endpoints should be required before the TST methods are placed into widespread use. If certain combinations of species/endpoints are found to have acceptable rates of false toxicity, it may be appropriate to use these species/endpoints with the TST method to trigger additional evaluations and study in the context of a narrative toxicity objective.

[10] SWRCB, 2004. Water Quality Control Policy for Developing California's Clean Water Act 303(d) List. September.

Table 2. Summaries of *Ceriodaphnia dubia* reproduction "blank" data from the USEPA Inter-Laboratory Validation Study. (Samples that were determined invalid by USEPA were not included).

Row #	Sample ID	Analysis Using the Proposed New TST Method					Current 40 CFR 136 Method	
		Mean Control Response	Mean Sample Response	% Effect	TST Results	Discharger has Reasonable Potential (RP) according to Draft Policy for Toxicity Assessment and Control	NOEC	IC 25
1	9330	25.4	25.0	1.5	Non-Toxic	No	100	>100
2	9332	16.6	16.3	1.8	Non-Toxic	No	100	>100
3	9337	20.1	19.4	3.5	Non-Toxic	No	100	>100
4	9338	24.2	21.3	12.0	Non-Toxic	**Yes**	100	>100
5	9340	15.3	19.8	-29.4	Non-Toxic	No	100	>100
6	9341	23.5	21.3	9.4	Non-Toxic	No	100	>100
7	9344	11.1	17.0	-53.2	Non-Toxic	No	100	>100
8	9349	30.8	30.3	1.6	Non-Toxic	No	100	>100
9	9350	29.5	22.9	22.4	**Toxic**	**Yes**	100	>100
10	9356	24.1	22.4	7.1	Non-Toxic	No	100	>100
11	9367	22.2	16.7	24.8	Non-Toxic	**Yes**	100	>100
12	9371	19.9	21.3	-7.0	Non-Toxic	No	100	>100
13	9376	20.4	17.8	12.7	Non-Toxic	**Yes**	100	>100
14	9379	24.9	26.8	-7.6	Non-Toxic	No	100	>100
15	9381	26.5	25.6	3.4	Non-Toxic	No	100	>100
16	9382	26.1	25.7	1.5	Non-Toxic	No	100	>100
17	9384	15.5	18.7	-20.6	Non-Toxic	No	100	>100
18	9402	16.0	16.2	-1.3	Non-Toxic	No	100	>100
19	9409	22.2	26.3	-18.6	Non-Toxic	No	100	>100
20	9410	24.8	22.8	8.1	Non-Toxic	No	100	>100
21	9429	31.0	31.1	-0.3	Non-Toxic	No	100	>100
22	9432	17.0	18.2	-7.1	Non-Toxic	No	100	>100
23	9436	28.1	31.8	-13.2	Non-Toxic	No	100	>100
24	9439	18.9	12.1	36.0	**Toxic**	**Yes**	100	>100
25	9445	23.6	22.4	5.1	Non-Toxic	No	100	>100
26	9446	22.2	18.3	17.6	**Toxic**	**Yes**	100	>100
27	9450	19.4	4.1	78.9	**Toxic**	**Yes**	25	15.9
Summary Statistics	N	27	27	27			27	27
	Min	11.10	4.10	-53.15			25	15.9
	Max	31.00	31.80	78.87			100	>100
	Median	22.20	21.30	1.81			100	>100
	Mean	22.20	21.17	3.29				>100
	# of Blank Samples Incorrectly Declared Toxic or Triggering Reasonable Potential				**4**	**7**	**1**	**1**
	Error Rate for Non-Toxic Blank Samples				**14.8**	**25.9**	**3.7**	**3.7**

Table 3. Probability of false violation errors with the assumed 15% error rate.

No. of false violation errors observed in 60 statistical analyses	Probability of observing **exactly** as many false violation errors	Probability of observing **at least** as many false violation errors
0 in 60	0.006%	-
1 in 60	0.06%	99.99%
2 in 60	0.3%	99.93%
3 in 60	1.1%	99.61%
4 in 60	2.8%	98.52%
5 in 60	5.4%	95.76%
6 in 60	8.8%	90.32%
7 in 60	12.0%	81.52%
8 in 60	14.0%	69.53%
9 in 60	14.3%	55.52%

C. Rates of toxicity for *C. dubia* in "real world" datasets are similar to those for USEPA blank data.

Table 4 presents a summary of the application of the TST method (*C. dubia*, reproduction endpoint) to stormwater samples collected by the County of Los Angeles. This dataset included 123 receiving water samples collected between 2005 and 2010. As shown in Table 4, the TST method found apparent toxicity in 12% of samples, and the methods of the Draft Policy would have led to findings of reasonable potential in 15% of the samples in this dataset. These rates of toxicity are comparable to the rates of toxicity in the USEPA blank dataset (see Section B, above) and are higher than the rates of toxicity found by the NOEC and IC25 methods (4% and 7%, respectively) when applied to the same dataset. These data indicate that it is likely that the rate of apparent toxicity of the TST method as applied to "real world" samples will be similar to the rate of apparent toxicity in the USEPA non-toxic blank samples, and higher than the rates of toxicity in the NOEC and IC25 methods approved by USEPA.

These data highlight the need to evaluate the TST method for various species and endpoints in both ambient samples and non-toxic blank samples. Because data for other species and endpoints in ambient samples are relatively scarce, evaluating the TST method may require the collection of additional data for a range of conditions (including stormwater and dry weather conditions), species, and endpoints, and evaluating the relationship between the sublethal methods and environmental effects (see also the discussion in Sections D and E below). Applying the TST methods before the evaluation of additional data would be premature.

Table 4. Los Angeles County Stormwater WET data of *Ceriodaphnia dubia* Chronic reproduction toxicity from 2005-2010.

Sample ID	Analysis Using the Proposed New TST Method					Current 40 CFR 136 Method	
	Mean Control Response	Mean Sample Response	% Effect	TST Results	Discharger has Reasonable Potential (RP) according to Draft Policy for Toxicity Assessment and Control	NOEC	IC25
# of Samples Declared Toxic				**15**	**18**	**5**	**9**
Rate for Toxic Samples				**12**	**15**	**4**	**7**

D. Sublethal endpoints for chronic toxicity testing are not supported by science

The proposed Draft Policy requires chronic sublethal toxicity tests (e.g., reproduction or growth). The chronic sublethal endpoints are much less reliable than a lethal endpoint (i.e., survival). Because chronic sublethal endpoint tests are less reliable than other test endpoints, they may not indicate any impact in ambient waters, and it would be inappropriate to determine noncompliance based on the outcome from a single sublethal toxicity test.

USEPA's own experts also acknowledge that effluent toxicity test failures based on the sublethal chronic toxicity tests using *C. dubia* reproduction or Fathead minnow growth or algal cell density should be considered suspect:

> *"…we continue to struggle with the idea that sublethal effects on indicator species can result in detectable adverse ecosystem response."*[11]

Sublethal endpoints are an unreliable indicator of actual instream impairment unless the reduction in growth or reproduction is caused by premature death in the test organisms. USEPA never conducted any scientific studies that would support correlation between sublethal endpoints and instream conditions.

EPA's conclusions are consistent with independent scientific research on the predictive power of chronic toxicity testing. For example:

> *"There is nearly a 50% probability that toxicity exhibited in WET tests may not be reflected instream, even for those effluents exhibiting a relatively high failure rate (>90%)*

[11] USEPA. A Review of Single Species Toxicity Tests: Are the Tests Reliable Predictors of Aquatic Ecosystem Community Responses? USEPA/600/R-97/114. July, 1999 at p. 24.

… A surprising result of this study was <u>the lack of relationship between Ceriodaphnia acute or chronic WET endpoints and instream biological results</u>."[12] (emphasis added)

"WET tests are not always reliable predictors of receiving environment conditions … Alone, WET tests cannot fulfill their stated purpose ("to identify, characterize, and eliminate toxic effects of discharges on aquatic resources")."[13]

Other states, such as Texas and Colorado, do not use growth or reproduction results to assess compliance with whole effluent toxicity limitations in NPDES permits. USEPA explains:

"EPA has not mandated which test methods NPDES permitting authorities must use under different exposure conditions…The permit writer has considerable discretion in selecting the appropriate test method (i.e., which test) as long as the method selected is consistent with the State's water quality standards and will protect the individual water in question, including the designated use…The permit writer must exercise his or her best professional judgment pursuant to 40 CFR 122.44(d), or analogous State law, to derive an appropriate WET limit, taking into consideration any State policies and or procedures for interpreting the narrative criteria and any available site-specific information."[14]

"The interpretation and application of [toxicity] test results are part of the implementation policy and are not addressed in this rulemaking [promulgating WET test methods] ..."[15]

Growth and reproduction data may not provide an accurate indication of the potential for toxicity to occur in the receiving waters. As detailed above, ionic interference is one reason the sublethal tests provide less reliable information.

Given USEPA's admitted uncertainty over the utility of sublethal endpoints, it is incumbent upon the State Board to provide more substantial scientific support for the necessity of these endpoints particularly in light of the ionic interference problem.

When a regulatory agency intends to use a method to evaluate compliance with the Clean Water Act, the highest level of validation is required:

"…methods which will be used extensively for regulatory purposes or where significant decisions must be based on the quality of the analytical data normally require <u>more</u>

---

[12] Diamond, J. and C. Daley. 2000. What is the relationship between whole effluent toxicity and instream biological condition? Environ. Toxicol. Chem. 19:158-168

[13] Chapman, P.M. 2000. Whole effluent toxicity testing-usefulness, level of protection, and risk assessment. Environ. Toxicol. Chem. 19:3-13.

[14] USEPA. "Clarifications Regarding Whole Effluent Toxicity Test Methods Recently Published at 40 CFR 135 and Guidance on Implementation of Whole Effluent Toxicity in Permits." Memorandum from Tudor Davies, Director Office of Science and Technology to Water Management Division Directors and Environmental Services Division Directors, Regions I-X. July 21, 1997 at pp. 2 & 3.

[15] USEPA. Whole Effluent Toxicity: Guidelines Establishing Test Procedures for the Analysis of Pollutants - Supplementary Information Document (SID) October 2, 1995 at pp. 28 & 33

extensive validation and standardization than methods developed to collect preliminary baseline data."[16] (emphasis added)

If sublethal test results are to be used for important regulatory purposes, such as discharge permitting and water quality monitoring, then the State Board must validate the methods for those uses. USEPA acknowledged that its Part 136 rule:

> *"…does not specify means to adjust [WET test methods] for the frequency, duration, or magnitude of instream exposure conditions…regulators need to make those "adjustments" when setting water quality standards and making permitting decisions."*[17]

Therefore, the State Board must discourage the use of sublethal endpoints as primary indicators of toxicity until these endpoints have been validated. Validation should involve demonstrating instream impacts even when toxicity tests are not failing the lethality endpoints (i.e., survival).

E. No correlation between the test results and instream condition was proven

Neither USEPA nor the State Board has conducted any studies to show the TST results are correlated with instream conditions. The proposed new toxicity objectives must be reasonably determine whether a discharge could affect biological integrity of aquatic environment. The legal standard for demonstrating that effluent toxicity testing should be included in a NPDES permit was given by one of USEPA's own administrative law judges:

> *"There must be a reasonable basis to believe the permittee discharge could be or become acutely toxic. In addition, the proposed [WET] tests must be reasonably related to determining whether the discharge could lead to 'real world' effects. The Clean Water Act objective to prohibit the discharge of 'toxic pollutants in toxic amounts' concerns toxicity in the receiving waters of the United States, not the laboratory tank"*[18]

An independent investigation initiated by the Water Environment Research Foundation with grant support by USEPA demonstrates that whole effluent toxicity test results are a poor predictor of ecosystem integrity.[19]

In addition, sublethal chronic toxicity endpoints were never examined for correlation with instream conditions by USEPA. USEPA's generic stream studies examined only the lethal endpoint (i.e., survival) for the chronic toxicity test and never examined the sublethal endpoints (i.e., reproduction and growth) independently.[20,21]

---

[16] Availability, Adequacy, and Comparability of Testing Procedures for the Analysis of Pollutants Established Under Section 304(h) of the Federal Water Pollution Control Act - Report to Congress; EPA/600/9-87/030; September, 1988; p. 3-5

[17] Settlement Agreement, *Edison Electric Institute, et al. v. EPA*, No. 96-1062 and consolidated cases (D.C. Cir.) (July 24, 1998)

[18] Andrew S. Pearlstein. In the Matter of Metropolitan Dade County (Fla.), Miami-Dade Water and Sewer Authority (NPDES Permit No. FL02241005), 1996 USEPA ALJ Lexis 100 (Oct. 3, 1996). Also cited in Water Environment and Technology, May 1997, p. 104).

[19] Water Environment Research Foundation. Evaluating Whole Effluent Toxicity as an Indicator of Instream Biological Conditions. Project 95-HHE-1. 1999.

[20] USEPA (1991). Technical Support Document For Water Quality-based Toxics Control. EPA/505/2-90-001. Washington, DC, Office of Water. p. 7; Hall, L.W. and J.M. Giddings. The Need for Multiple Lines of Evidence for Predicting Site-Specific

The proposed Draft Policy will apply the numeric toxicity objectives to waterbodies including ephemeral or effluent-dependent waterbodies, which are common in California. There is very little evidence to indicate that the chronic toxicity test methods for both lethal and sublethal endpoints provide a meaningful measure of biological integrity in ephemeral or effluent-dependent ecosystems.

USEPA stated in response to a Freedom-of-Information Act request that it has "no information" to demonstrate that effluent whole toxicity test endpoints are correlated with biological conditions in effluent-dominated streams, stormwater channels or agricultural drains.[22] USEPA field studies did not validate the use of chronic toxicity test results in effluent-dominated streams. In fact, no streams west of Texas and Oklahoma were included in USEPA's Complex Effluent Toxicity Testing Program (CETTP) studies[23] that were designed to validate the use of whole effluent toxicity test methods for NPDES permitting purposes. In addition, the CETTP studies did not investigate the statistical relationship between effluent toxicity and biological community effects in the receiving water[24].

F. No site-specific consideration for California waterbodies

Changes in natural water chemistry can interfere with test results. The Draft Policy includes no guidance regarding how to account for the resulting artifacts when interpreting results. The natural ionic chemistry of California receiving waters and local groundwater supplies may interfere with the normal growth and reproduction of test organisms. This in turn may produce erroneous indications of toxicity in the chronic test method, particularly in the sublethal endpoints.

USEPA guidance warns that small changes in the ionic balance may interfere with the conduct and interpretation of effluent toxicity tests using standard indicator organisms.

> *"Hardness of dilution water, as well as hardness of water used to culture test organisms, may have an effect on successful completion of the tests. In some cases, the relative hardness of the dilution water compared to the organism culture water may affect the*

Ecological Effects. Human and Ecologial Risk Assessment. Vol. 6, No. 4. Pp. 678-710 (2000); Parkhurst, B. "Predicting Receiving System Impacts from Effluent Toxicity." Whole Effluent Toxicity Testing: An Evaluation of Methods and Prediction of Receiving System Impacts. Society of Environmental Toxicology and Chemistry. Proceedings of the Pellston Conference. 1996; Diamond, J. and C. Daley. "What is the relationship between whole effluent toxicity and instream biological condition?" Environ. Toxicol. Chem. 19:158-168 (2000).

[21] WERF conducted a pilot study to evaluate the use of WET testing in effluent dominated streams (see 'Evaluation of WET testing as an indicator of aquatic health in effluent-dominated streams: a pilot study (WERF 2009)). In the pilot study, effluent samples were collected only from six facilities in five different states and were tested for both lethal and sublethal endpoints in the chronic WET tests. However, none of the effluent samples were from states in the west of Colorado. This was not a full scale study but a pilot study, and the number of samples included in the study was too small to draw any conclusion. Based on the limited data, authors of the study concluded that " WET test results exhibited few relationships with bioassessment results and could not usually predict instream effects even when incorporating actual effluent dilution." (p. V)

[22] USEPA. Response to Freedom of Information Act request (submitted 5/28/96 and resubmitted 7/24/96). Letter to Mark Pifher, attorney of record for Western Coalition of Arid States. September 11, 1996.

[23] The eight CETTP studies include Scippo Creek, Ohio, Ottawa River, Ohio, Skeleton Creek, Oklahoma, Fiver Mile Creek, Alabama, Ohio River, West Virginia, Kanawha River, West Virginia, Naugatuck River, Connecticut, and Back River, Baltimore Harbor, Maryland. (See Parkhurst, B. 1994. Chapter 7 Are single species toxicity test results valid indicators of effects to aquatic communities? in Ecological Toxicity Testing: Scale, Complexity, and Relevance. Cairns Jr., J. and Niederlehner, B., CRC Press)

[24] Parkhurst, B., Marcus, M. and Noel, L. (1990). Review of results of USEPA's Complex Effluent Toxicity Testing Program. Washington, DC, Utility Water Act Group: 181.

expression of toxicity in the conduct of the tests, i.e., the accuracy of the tests at predicting toxicity." [25]

In fact, during the large scale whole effluent toxicity variability study, USEPA disqualified test results when hardness concentrations in the control water varied by only 15-20% from the nominal values.[26] The results were rejected when it was feared that ionic interference may result in false indications of toxicity. In Southern California, the hardness of effluent samples and receiving waters are routinely much higher than the control water. Consequently, there is a serious risk of bias when standard test species are used to evaluate the potential for toxicity in these waters.

"In addition, parameters such as TDS (hardness, salinity, conductivity), turbidity, DO, pH, micronutrients, and bacteria counts can impact test organism physiology, sensitivity, and biological response. Therefore, test variability at all levels can be affected by variability in dilution water quality." [27]

USEPA guidance is consistent with other peer-reviewed scientific studies. According to Goodfellow et al. (2000), even small differences in the balance between major anions and cations can interfere with the normal reproductive cycle of some standard WET test species.[28] Chapman (2000) argues that such interference can result in artificial failures in the WET test:

"…test organisms may be sensitive to noncontaminant effects. For instance, increased hardness is a feature of some effluents, which can have an adverse effect on daphnids irrespective of contaminant concentrations. Variations in salinity and total dissolved solids can significantly affect WET test organisms." [29] (emphasis added)

Proof of this ionic interference problem was recently developed by the Western Coalition of Arid States and submitted to USEPA for their consideration.[30] Results from that study clearly demonstrate that the average ionic composition of western waters is sub-optimum for supporting survival and reproduction in *C. dubia* (an invertebrate species found largely in lakes and ponds of relatively low hardness in the northern United States).

Even if one assumes that testing is an appropriate regulatory tool to assess eastern rivers and streams, the absence of validation for common conditions in the arid west would be a significant error.

---

[25] USEPA. Understanding and Accounting for Method Variability in Whole Effluent Toxicity Applications Under the NPDES Program. EPA-833-R-00-003 (June, 2000); at p. D-6 & 7

[26] USEPA. Final Report: Interlaboratory Variability Study of EPA Short-term Chronic and Acute Whole Effluent Toxicity Test Methods-Vol. 1 & 2; EPA-821-B-01-004; September, 2001

[27] Understanding and Accounting for Method Variability in Whole Effluent Toxicity Applications under the NPDES Program. EPA-833-R-00-003 (June, 2000); p. D-7 (WET-IX Docket #B.12)

[28] Goodfellow, W.L., et al. 2000. Major ion toxicity in effluents: a review with permitting recommendations. Environ. Toxicol. Chem. 19:175-82.

[29] Chapman, P.M. 2000. Whole effluent toxicity testing-usefulness, level of protection, and risk assessment. Environ. Toxicol. Chem. 19:3-13.

[30] Declaration of Timothy F. Moore and accompanying memorandum from Mr. Moore to John Hall, Attorney at Law, dated July 25, 2004 and June 23, 2004 respectively. Affidavit submitted to U.S. Court of Appeals - D.C. Circuit in Case No. 96-1062.

"General validation consists of testing, evaluating and characterizing the method to the extent necessary to demonstrate that the method achieves a specified performance. This process establishes quantitative measures of performance under typical conditions of use … A method that is generally validated cannot unequivocally be assumed to be valid for every specific use." [31]

G. The proposed Draft Policy fails to consider significant sources of experimental test bias

As discussed previously, toxicity testing results can be affected by factors other than actual toxicity of effluent samples. However, the Draft Policy fails to consider the variability of testing results due to these factors and provides no guidance to resolve these issues.

USEPA points to ambient hardness as an example of the sort of factors that may increase problems with analytical variability in effluent toxicity testing:

"Abiotic conditions can strongly influence the variability of WET test results. For that reason, most of the abiotic conditions that should be standardized during WET testing (DO, light, hardness, alkalinity, etc.) are specified in the USEPA method manuals. While these factors may not be problematic sources of variability within tests, they may be of major concern across tests (both within and among laboratories). …Careful use of dilution waters, salinity adjustments, aeration, feeding, and other factors causing shifts in pH will help reduce variability." [32]

Independent, peer-reviewed scientific studies have shown that more than half of the observed variations in whole effluent toxicity test results are due to factors other than actual effluent quality.[33] And, USEPA has acknowledged that the relative error band for most toxicity tests is "plus or minus 100%."[34]

H. The use of surrogate species is inappropriate to assess a permit violation.

Standard species used in toxicity tests may or may not be relevant to the specific aquatic ecosystem that whole effluent toxicity tests are intended to protect. As Chapman notes, it is inappropriate to use surrogate species as an evidence of a permit violation for toxicity.

"Whole effluent toxicity test species are generally not the same as resident species that the results of WET testing are aimed at protecting … Differences exist between sensitivities and tolerances of WET species. Such differences are not unexpected … however, these differences can become profound when regulatory use of WET test results involves a bright line that does not adequately account for species differences." [35]

---

[31] USEPA, *Availability, Adequacy, and Comparability of Testing Procedures for the Analysis of Pollutants Established Under Section 304(h) of the Federal Water Pollution Control Act, Report to Congress*,EPA/600/9-87/030 (September 1988) at pg.3-6.

[32] Understanding and Accounting for Method Variability in Whole Effluent Toxicity Applications Under the NPDES Program. EPA-833-R-00-003 (June, 2000); p. D-6 (WET-IX Docket #B.12)

[33] Water Environment Research Foundation. Whole Effluent Toxicity Variance Report. 1999 at pg. 3-3.

[34] USEPA. Short-Term Methods for Estimating Chronic Toxicity of Effluents and Receiving Water to Freshwater Organisms - 4th Ed.; EPA-821-R-02-013; October, 2002. See section 4.14.6

[35] Chapman, P.M. 2000. Whole effluent toxicity testing-usefulness, level of protection, and risk assessment. Environ. Toxicol. Chem. 19:3-13.

I. The use of *C. dubia* may lead to unreliable outcomes of toxicity testing.

The proposed Draft Policy will require the use of the most sensitive species for the routine compliance monitoring. It is highly likely that *C. dubia* will be selected as the species because *C. dubia* in general fails more frequently in toxicity testing than other test species (e.g., fathead minnow or green algae), and this high failure rate is deemed by the Draft Policy to indicate high sensitivity. However, higher failure rates could imply that *C. dubia* might not be a reliable test organism (i.e., one which yields reproducible and consistent results during toxicity testing). The Draft Policy includes no guidance to determine the difference between the sensitivity and the unreliability of species in the selection of the most sensitive species.

C. dubia are hyper-optimized to laboratory test water, which differs significantly from waters in the environment. For this reason, *C. dubia* become overly sensitive to any differences in the composition of test waters as compared to laboratory test water (e.g., changes in ion strength, presence of chlorine). *C. dubia* are cultured in a freshwater matrix that more closely resembles the low-hardness water from Lake Michigan than the ground and surface waters common to many regions in California which exhibit higher hardness and higher TDS concentrations. The same Great Lakes recipe for laboratory water is used as the default control water in toxicity tests, thereby biasing the result of WET tests in California.

On average, *C. dubia* exposed to synthetic western groundwater produced 30% fewer offspring in the allotted 7-day test period than organisms assigned to a control group.[36] Elevated hardness does not so much preclude reproduction as it temporarily defers reproduction until the test organisms acclimate to the new ionic matrix. However, in most instances the toxicity test is terminated before the effluent-exposed organisms are able to acclimate and "catch up" to the control group. Therefore, according to USEPA, the test failure is not always a valid indication of actual instream toxicity:

> *"It is not always obvious that an effect level that is determined to be statistically significant is also biologically significant. The implied question, concerning the biological significance of (threshold) statistically significant occurrences of adverse biological effects observed in toxicity tests, is an implementation question, and is not addressed in this rulemaking."*[37]

The State Board should consider the use of other surrogate organisms[38] or dual control techniques recommended by USEPA[39] in order to reduce the experimental test bias.

---

[36] Declaration of Timothy F. Moore and accompanying memorandum from Mr. Moore to John Hall, Attorney at Law, dated July 25, 2004 and June 23, 2004 respectively. Affidavit submitted to U.S. Court of Appeals - D.C. Circuit in Case No. 96-1062.

[37] USEPA. Whole Effluent Toxicity: Guidelines Establishing Test Procedures for the Analysis of Pollutants - Supplementary Information Document (SID). Oct. 2, 1995. p. 33

[38] See, for example, the San Diego Regional Board's approval of Hyallela azteca for WET tests assessing groundwater samples with naturally elevated TDS concentrations. California Regional Water Quality Control Board-San Diego Region. "Revised Requirements for Enrollment Under Order No. 2001-96; NPDES No. CAG919002. August 19, 2004. (Letter to Kevin J. Ryan, Kinder-Morgan Energy Partners).

[39] Method Guidance and Recommendations for Whole Effluent Toxicity (WET) Testing (40 CFR Part 136). EPA-821-B-00-004; July, 2000 p. 6-5 – 6-6. "For each test using receiving water for dilution, a 100% receiving water control and a 100% culture water control should be run concurrently in the test and compared to determine the presence of toxicity in the receiving water."

J. The proposed Draft Policy precludes confirmation of a valid dose-response relationship.

A valid dose-response is essential to confirm the present of toxicity, and multiple concentration tests are required to evaluate the dose-response. The simple, two-concentration tests (i.e., a control of zero concentration of an effluent sample and a treatment of a 100% concentration of an effluent sample) specified in the Draft Policy are inadequate for this purpose. In fact, determining compliance using one instream waste concentration (IWC) does not meet the minimum number of four samples and of two species required under Title 40 § 122.21[40].

K. The TST method is not approved for use under current legal requirements.

40 CFR Part 136 contains guidelines establishing test procedures for the analysis of pollutants. The TST method is new (released by USEPA in June 2010), and was never approved as meeting these guidelines. No federal register notices were released concerning the TST guidance, and public comments were not requested for the TST method. No peer review that meets California requirements was conducted by USEPA. No estimates were made of inter-laboratory test precision, which is required for any new method. It has not yet been authorized as an alternate test procedure pursuant to 40CFR 136.5 approval of alternate test procedures. For these reasons as well as for the technical reasons detailed in this comment letter, use of the TST method is premature and unfounded.

In addition, the requirements appear to be more stringent than necessary to comply with the Clean Water Act as the Draft Policy requires the use of numeric effluent limitations when they are not otherwise required by federal or state law.[41] In addition, the Draft Policy is more stringent than federal law as it automatically assumes reasonable potential for a certain group of dischargers (i.e., those discharging more than 1 mgd), which is contrary to the federal rule [40 C.F.R. §122.44(d)(1)(iv)][42] that prescribes an effluent limit for toxicity only where a discharge has "the reasonable potential to cause, or contributes to an in-stream excursion above the numeric criterion for whole effluent toxicity." Thus, it appears that the adoption of the TST method as State Board policy would create an unfunded mandate. Most of all, federal law and regulation do not require the use of the TST method. If the State enacts the TST method, it will lead the state to deviate from 40 CFR Part 136 methods.

L. The WET policy is unnecessarily stringent.

The proposed TST method is significantly more stringent than existing, USEPA-approved procedures for toxicity determination (e.g., the IC25 or NOEC). Presently, if effluent-exposed organisms exhibit at least 75% of the reproduction or growth shown by control organisms, the effluent is presumed non-toxic. The TST method will require that effluent-exposed organisms exhibit at least 90% of the growth or reproduction shown by the control group[43]. As shown in Tables 2 and 4, the TST method (for

[40] 40 CF 122.21(g)(7) …. the minimum of four (4) grab samples will be a representative sample of the effluent being discharged…. (j)(5) (v) Applicants must conduct tests with multiple species (no less than two species; e.g., fish, invertebrate, plant), and test for acute or chronic toxicity, depending on the range of receiving water dilution.

[41] Communities for a Better Environment (CBE) v. State Board/Tesoro,109 Cal.App.4th 1089, 1103–07 (2003); State Board Order No. WQ 91-03, 1991 WL 135460, at 12; 40 C.F.R. §122.44(k).

[42] (iv) When the permitting authority determines, using the procedures in paragraph (d)(1)(ii) of this section, that a discharge causes, has the reasonable potential to cause, or contributes to an in-stream excursion above the numeric criterion for whole effluent toxicity, the permit must contain effluent limits for whole effluent toxicity

[43] There is a discrepancy in the determination of toxicity between the USEPA's TST method and the proposed Draft Policy. USEPA's TST method specifies that an effluent sample will be determined as toxic if it fails in the TST test AND a mean percent effect is same as or larger than 10%. However, the Draft Policy will determine a sample as toxic if the TST test fails regardless of the mean percent effect. WSPA respectfully suggests that the determination of toxicity should be revised to

C. dubia, reproduction endpoint) will result in higher rates of apparent toxicity as compared to the IC25 and NOEC methods currently approved by USEPA.

While issuance of the USEPA TST Guidance means that there is now an additional statistical approach to determine toxicity in addition to the existing NOEC and EC/IC25 methods, the only statistical analysis specifically recommended for NPDES compliance purposes in the promulgated toxicity methods remains the point estimate technique (i.e., EC/IC25). Both the TST Guidance and the USEPA Technical Support Document for Water Quality Based Toxics Control concur that the appropriate effect level threshold for defining unacceptable chronic toxicity is 25% or an EC/IC25. As discussed above, WSPA requests that the State Board perform additional data evaluations using both blank and ambient samples for a range of species and endpoints. Only those measures with reasonable rates of false toxicity should be used in State Board policy, and they should be used as triggers for additional study and not as numeric effluent limitations.

3. The reasonable potential analysis (RPA) of the Draft Policy will result in unnecessary application of effluent limitations.

The Draft Policy results in a finding of reasonable potential (i.e., a finding that an effluent has the potential to cause or contribute to an excursion above a water quality standard, and thus requires an effluent limitation) under two conditions: (1) if an effluent sample fails the TST, or (2) if the percent effect (i.e., the difference between responses of the effluent sample and the control) is greater than 10%. The Draft Policy provides no scientific justification for the reasonable potential threshold of 10%. The 10% difference from control is unfounded and unreasonable, considering the variability inherent in the tests. The 10% threshold has no relationship to actual toxicity, and is not included in USEPA's TST methods.

As shown above in Table 2, the false failure rate of the RPA in the USEPA's blank data for *C. dubia* chronic toxicity testing is about 26% (i.e., seven of 27 samples). At this rate, a discharger conducting monthly monitoring for the purpose of assessing reasonable potential has only a 0.01% probability of passing the RPA during a 5-year permit term, even if effluent is as non-toxic as blank reference water. The failure of the RPA will result in the unnecessary application of effluent limitations.

4. The proposed Draft Policy will dramatically increase the use of animals in toxicity tests.

If a fish species is identified as the most sensitive species and the minimum of four replicates per toxicity test are used (i.e., a total of 8 individuals—four for controls and four for effluent samples) as shown in examples of the TST implementation document (USEPA 2010), a minimum of 96 fish per year will be required for monthly routine monitoring. If the acute numeric objectives as well as the chronic objectives are applied, the number of fish required could double to 192 per year. If the discharger failed a test, the accelerated monitoring requirements in the Draft Policy will increase the number of fish dramatically, resulting in the use of 144 fish per year at the USEPA-assumed false violation rate of 5%.

---

comport with USEPA's TST method because the Staff Report states that "[t]he proposed Policy does not depart from the scientific approach of the TST" (at p. 65).

If *C. dubia* is identified as the most sensitive species, 240 individuals will be required to conduct monthly routine monitoring per year; 10 replicates are required for *C. dubia* testing as shown in examples of the TST implementation document (USEPA 2010). If a discharger fails a test, this single-test failure will trigger accelerated monitoring and will require a minimum of 600 individuals. At a false violation rate of 15% (see Table 2), this will result in the use of 1,080 *C. dubia* organisms per year, even if an effluent is entirely non-toxic.

This excessive use of animals for toxicity tests is contrary to the global and regional efforts of WSPA member organizations to "reduce, refine, and replace" animal testing.

5. The proposed Draft Policy should be adopted through a formal rule-making process.

The State Board's definition of whole effluent toxicity using the entirely new TST method constitutes a change in water quality standards. Thus, it must be adopted through a formal rule-making process rather than as policy guidance, and the State Board must first comply with the requirements in California Water Code Sections 13241 and 13242.

6. The cost analysis in the Staff Report significantly underestimates costs and environmental impacts of the Draft Policy

The economic analysis contained in the Staff Report for the Draft Policy fails to consider the reasonably foreseeable cost of compliance. With regard to monitoring, monthly monitoring costs can increase from one species to three species in order to identify the most sensitive species as required. If a discharger's effluent samples are not found to have reasonable potential (i.e., they pass the RPA), the discharger will likely be required to conduct monthly RPA monitoring using three species. The cost savings that will result from testing only two concentrations, rather than 6 concentrations as required by current methods, will not be significant. But more significantly, the number of tests that will be performed under the accelerated monitoring requirements of the Draft Policy are expected to increase, as the rate at which non-toxic samples are erroneously determined to be toxic will increase significantly. In addition, the number of TIE/TRE studies that will be required pursuant to false findings of toxicity will increase significantly. Further, it appears that the Staff Report dramatically underestimates the cost of performing TIE/TRE analyses. For example, San Jacinto Water Authority (near Houston, TX) spent more than $250,000 on a TIE/TRE in 2003-2004 to address random and infrequent failures of the *C. dubia* reproduction test. One cause of the failures was determined to be test interference caused by the naturally-elevated conductivity, alkalinity and hardness of local water supplies.[44]

[44] More examples for the costs of TIE/TREs are provided below;

- Inland Empire Utilities Agency (Chino, CA) spent more than $300,000 on a TIE/TRE in 1997-98 to address sporadic failures of the *C. dubia* reproduction test.
- Georgia-Pacific (Palatka, FL) has spent nearly $500,000 on various TIE/TRE studies from 2003-2009 to address failures of the *C. dubia* reproduction test. This appears to be a case where the alum used in the water treatment process is interacting with the unusual chemistry of local groundwater to create an effluent that is toxic in the laboratory but not in the field. Florida and EPA are still working to determine how to address this situation.
- City of San Bernardino spent more than $100,000 on accelerated tests and preliminary TIEs over the last 10 years. In every instance, it appears that the initial failure of the *C. dubia* reproduction test was due to routine (annual) culture crashes at the analytical laboratory.
- Milliken and Co., Inc. (SC) operates numerous textile mills in South Carolina and has spent nearly $1 million on TIE/TRE studies at 5-6 different mills over the last 10 years.

In addition to monitoring costs, the Draft Policy would require significant expenditures to achieve compliance with the requirements of the TST policy. For example, it could be reasonably anticipated that full nitrification may be required to eliminate ammonia in discharges regulated by the policy. Disinfection by ultraviolet light (UV) or ozone may be indicated to reduce toxicity from chlorination and dechlorination processes. Activated carbon may be required to reduce concentrations of organic compounds. Reverse osmosis (RO) may have to be implemented to reduced hardness and alkalinity and concentrations of total dissolved solids (TDS), which can contribute to toxicity test failures. All of these treatment methods are expensive, consume significant amounts of energy, generate greenhouse gases, and none of them were included in the cost analysis.

As noted throughout these comments, these measures may be indicated and implemented to remedy findings of toxicity result from the high rate of false toxicity that is expected to result from implementation of the TST method of the Draft Policy. Thus, the Draft Policy may result in implementation of these measures even in discharges that are not toxic, thus resulting in no environmental benefit.

The State Board staff claims "no adverse environmental impact" of any kind from the Draft Policy[45]. But at the same time, the Staff Report states that the staff was unable to assess any of the reasonably foreseeable compliance alternatives, as required by CEQA, because doing so would be "purely speculative." As detailed above, the Draft Policy is highly likely to result in numerous adverse environmental impacts. Construction of treatment systems will result in construction, noise, and traffic impacts. Operation of treatment systems will require substantial energy resources, will generate greenhouse gases, and will require disposal of waste streams. These and other reasonably foreseeable environmental impacts must be fully analyzed as required by CEQA prior to adoption of the Draft Policy.

7. Expansion of whole effluent toxicity testing to stormwater discharges is invalid.

The proposed Policy expands whole effluent toxicity testing to stormwater discharges (and discharges covered by agricultural waivers) and lacks appropriate studies or data collection to support this expansion. This will lead to a significant increase in enforcement actions and related appeals.

The Draft Policy appears to lack any analysis or consideration of the unique characteristics of stormwater, which differs significantly from point-source discharge. Stormwater discharges in California are intermittent, infrequent (especially in the southern part of the State), and exhibit highly variable flow rates and constituent concentrations. The treatment methods that may be applied to stormwater discharges differ in significant ways from those applicable to traditional point source discharges. As discussed previously, USEPA never validated whole effluent toxicity testing methods in stormwater channels, ephemeral stream or agricultural drains.[46]

- Springs Industries (GA) spent more than a $100,000 on a TIE/TRE at one textile finishing plant in Griffin, GA in 2001-2003.
- Kinder-Morgan has spent more than $100,000 on TIE/TRE work at a groundwater remediation project in San Diego. The final effluent passes through two activated carbon filters before being discharged. Ionic interference caused by the natural chemistry of the groundwater is the problem.

[45] State Water Resources Control Board, 2010. Staff Report for "Draft Policy for Toxicity Assessment and Control" at p. 62.

[46] USEPA. Response to Freedom of Information Act request (submitted 5/28/96 and resubmitted 7/24/96). Letter to Mark Pifher, attorney of record for Western Coalition of Arid States. September 11, 1996.

The State Board has provided no evidence or data to support the application of the TST approach to stormwater. As a result, the proposed Draft Policy provides no guidance for dischargers on how to conduct chronic sublethal toxicity tests in stormwater samples. Chronic sublethal toxicity tests require changing water in a treatment train every day for about eight days with a new effluent sample; however stormwater discharges frequently last only several hours or days. It is infeasible to collect stormwater samples every day for the period of the chronic test, as rain events typically do not last long enough. Even when stormwater flows persist for time periods of a week or more, the composition of stormwater samples will differ significantly.

Given the absence of general validation studies for effluent-dominated streams, stormwater channels and agricultural drains, the Draft Policy should not be used to determine reasonable potential or to establish effluent limitations for these types of discharges.

Recommendation

WSPA strongly recommends that the State Board not adopt the numeric objectives and the TST method of the Draft Policy. An approach of narrative objectives with accelerated monitoring and toxicity reduction evaluation (TRE) triggers is supported by current analyses and data, and would be appropriate for addressing effluent toxicity. This approach has been effectively implemented in California for several years, is consistent with guidance from USEPA,[47] and is supported by recognized national and regional experts[48].

The State Board should discourage the use of sublethal endpoints as primary indicators of toxicity until these endpoints have been validated. The State Board should evaluate a range of species and endpoints in both blank and ambient samples in order to assess the rate of false toxicity. For those species and endpoints that exhibit reasonable rates of false toxicity, the TST method could be used as a trigger for additional study within the framework of a narrative toxicity objective. The State Board should also consider using other surrogate organisms or dual control techniques as recommended by USEPA in order to reduce the experimental test bias.

[47] 'Technical Support Document for Water Quality-Based Toxics Control, USEPA Office of Water, March 1991, USEPA/505/2-90-001, p. 62, Section 3.3.7' and 'EPA Regions 9 and 10 Guidance for Implementing Whole Effluent Toxicity Testing Programs, USEPA, May 31, 1996, pp. 2-1, 4-1, and 5-2.'

[48] 'Society of Environmental Toxicology and Chemistry (SETAC) WET Expert Advisory Panels, http://www.setac.org/wettre.html, Sections 1 and 4.' and 'Memo to Members of the State Water Resources Control Board from the Toxicity Task Force, September 27, 1995. Recommendations 2, 5, 9, and 10.'