CENTRAL VALLEY CLEAN WATER ASSOCIATION
CVCWA
EST. 1987

Central Valley Clean Water Association

Representing Over Fifty Wastewater Agencies

TERRIE MITCHELL – Chair, Sacramento Regional CSD | **CASEY WICHERT – Vice-Chair, City of Brentwood**
JOSIE TELLERS – Secretary, City of Davis | **KEN GLOTZBACH – Treasurer, City of Roseville**

December 21, 2018

Jeanine Townsend
Clerk to the Board
State Water Resources Control Board
1001 I Street, 24th Floor
Sacramento, CA 95814
commentletters@waterboards.ca.gov

Re: Comment Letter — Toxicity Provisions

Dear Ms. Townsend:

The Central Valley Clean Water Association (CVCWA) appreciates the opportunity to comment on the Draft Water Quality Control Plan for Inland Surface Waters, Enclosed Bays, and Estuaries of California; Toxicity Provisions (Toxicity Provisions). CVCWA is a non-profit association of public agencies located within the Central Valley region that provide wastewater collection, treatment, and water recycling services to millions of Central Valley residents and businesses. We approach these matters with the perspective of balancing environmental and economic interests consistent with state and federal law.

First, we appreciate the time extension the Board has provided for this Comment Letter for CVCWA to finalize our Phase I Report for our Toxicity Special Study (Toxicity Report). This report was developed in collaboration with the Central Valley Regional Water Quality Control Board (Central Valley Water Board) and some of the Publicly Owned Treatment Works (POTWs) in the Central Valley to assess low level indications of toxicity and better understand whole effluent toxicity testing and its nexus to our waterbodies. The Toxicity Report is attached.

In this letter, we provide significant and other comments regarding the proposed Toxicity Provisions. As a preliminary matter, we want to convey our appreciation to the State Water Resources Control Board (State Water Board) for including provisions for insignificant discharges, as well as small disadvantaged communities.

In addition to the comments provided in this letter, CVCWA joins the comments made by the California Association of Sanitation Agencies (CASA) and the Bay Area Clean Water Agencies (BACWA) on the Toxicity Provisions.

Introduction

In the Executive Summary of the Draft Staff Report for the proposed Toxicity Provisions, it is stated that:

> "Aquatic toxicity occurs when the effects of pollutants in surface water negatively impact aquatic life beneficial uses. When originating from an effluent, these effects are typically referred to as 'whole effluent toxicity' (WET)."

CVCWA believes that whole effluent toxicity is more properly defined as toxicity measured in an effluent sample which is used as a surrogate to estimate toxicity in receiving waters.

Further, we question the allegation that significant evidence exists to demonstrate that the ambient toxicity which has been observed in California waters "originates from effluent." The statewide ambient toxicity results summarized in Section 4.2 of the staff report indicate that pesticides are the primary source of observed toxicity in ambient waters in California. As indicated in the attached Toxicity Report, a summary of 35 Toxicity Reduction Evaluation (TRE) investigations performed in the Central Valley Region of California since 2011 show that pesticides are not an observed cause of WET. No linkage has been made in the staff report (or in any other documentation supporting the proposed Provisions) between WET results and ambient toxicity observations or impairments. We believe that this information indicates that POTW discharges in California do not pose a significant risk to ambient water quality or receiving water aquatic life uses.

We believe that this information supports our position that the proposed Provisions should not be unnecessarily conservative in the implementation of the proposed water quality objectives in National Pollutant Discharge Elimination System (NPDES) permits. Our comments below reflect reasonable requests to move policy implementation in this direction.

Significant Comments

1. Instream Waste Concentration (IWC) Determinations

WET measurements are an indirect indicator of the toxicity of effluent discharges to receiving waters. Effects measured in whole effluent may not necessarily translate to similar effects in ambient waters. The level of hazard associated with an effluent is significantly influenced by the dilution of the effluent in receiving waters. The CVCWA Toxicity Special Study: Phase 1 Report (December 2018) (Toxicity Report), attached hereto as Attachment A, summarizes information from toxicological literature that reinforce the concept that WET results best reflect ambient conditions downstream of an effluent discharge when dilution is properly taken into account. This finding has been well established since the early days of the WET requirements in the NPDES program. (See 40 CFR 122.44(d)(1)(ii); see also United States Environmental Protection Agency (USEPA) Technical Support Document for Water Quality-Based Toxics Control (USEPA TSD) (1991), Section 1.3.2, p. 7.) It has also been corroborated by information presented in Diamond and Daley (2000). (Diamond, J. and Daley, C. 2000. *What is the Relationship Between Whole Effluent Toxicity and Instream Biological Condition?* Environmental Toxicology and Chemistry, Vol. 19, No. 1, pp. 158-168.) Diamond and Daley cited work which found that basing WET compliance on average or actual stream flow conditions more efficiently predicted instream biological conditions than the use of the seven-day average condition expected to occur once in 10 years (7Q10).

In light of the above, we request that the IWC language in the proposed Provisions be modified to allow regional water quality control boards (Regional Water Boards) flexibility to establish an IWC based on actual in-stream conditions during discharge events and/or to establish an IWC that accounts for seasonality. CVCWA requests that Section IV.B.2.d be revised to read as follows (revision shown in italics):

> On page 20, last paragraph, last sentence,
>
> "The DILUTION RATIO shall be determined using the parameters specified in Table 3, *or, alternatively, shall be determined using a method approved by the Permitting Authority that accounts for dilution conditions occurring in the receiving water during the period of the toxicity test, including consideration of seasonality*."

We also request that the language in Table 3 on page 21 of the Toxicity Provisions be modified so that the averaging periods match the duration of chronic toxicity tests, reduce unnecessary conservatism, and create a more accurate assessment of effects in ambient waters. Specifically, we request that the title of Table 3 be changed to "Parameters for Calculating a Dilution Ratio, unless otherwise approved by the Permitting Authority". We also request the following changes to the column in Table 3 titled "Use the Discharge Effluent Flow Of:"

Next, we request that "Maximum daily flow (i.e. the maximum flow sample of all samples collected in a calendar day)[1] during period of discharge" for acute toxcitiy be replaced with "*Average daily flow (i.e. the average of all flow measurements in a calendar day) during period of discharge.*" We also request that "Four-day average of daily maximum flows (i.e. the average of daily maximums taken from the data set in four-day intervals) during period of discharge" for chronic toxicity be changed to "*Four-day average of all flows (i.e. the average of all flow measurements taken in four-day intervals) during the period of discharge.*"

In addition, the State Implementation Policy (SIP) (Page 1.4.D on page 13) includes the following language regarding the consideration of seasonal conditions in establishing effluent limits for, among other parameters, chronic toxicity.

> "In determining the appropriate available receiving water flow, the RWQCBs may take into account actual and seasonal variations of the receiving water and the effluent."

CVCWA requests that the above provision be added to Section IV.B.2.d of the proposed Toxicity Provisions.

2. Reasonable Potential Determinations

Federal NPDES regulations at 40 C.F.R. section 122.44(d)(1)(ii) require that effluent limits be established where it is determined that a discharge "causes, has the reasonable potential to cause, or contributes to an in-stream excursion above . . . water quality standard." As it has been applied in NPDES permits in California, and consistent with the approach documented in the USEPA TSD and the SIP, the water quality standard is used directly in the determination of reasonable potential. With specific reference to WET, the federal regulations specifically refer to "an in-stream excursion above the numeric criterion for whole effluent toxicity." (40 C.F.R. § 122.44(d)(1)(iv).)

The Toxicity Provisions propose to use a metric (where test results indicate a 10 percent effect or greater) to determine whether a discharge has reasonable potential for both chronic and acute toxicity, but that is not the water quality objective. This approach also does account for dilution in the receiving water. This results in reasonable potential determinations that are significantly more conservative than is necessary. The Staff Report does not offer an adequate rationale to justify this overly conservative and unconventional approach.

To maintain consistency with the conventional basis for reasonable potential determinations, it is requested that the proposed language be changed to define Reasonable Potential for chronic toxicity based on either: (1) a 25 percent effect; or (2) a

[1] Note that the definition of Maximum Daily Flow is inconsistent with prior definitions for Maximum Daily Flow. What is described in parenthesis is instantaneous maximum.

failed test at the IWC as determined using the TST. For acute toxicity, it is requested that the proposed language be changed to define Reasonable Potential to be based on either a (1) 20 percent effect or (2) a failed test at the IWC using the TST. The suggested change to the percent effect for chronic and acute toxicity are consistent the water quality objectives set forth in the Toxicity Provisions on page 2. These specific requested language changes are to replace "10 percent" with "*25 percent*." In the third paragraph, replace "10 percent" with "*20 percent*" in Section IV.B.2.b. on page 15, second paragraph.

Please see the following for a strikeout version of our proposed revisions to Page 15, Section 2.b.i:

> "A discharge has REASONABLE POTENTIAL to cause or contribute to an excursion above the chronic toxicity water quality objectives specified in Section III.B.2.a, if any of the CHRONIC TOXICITY TESTS result in a "fail" at the IWC, or if any of the CHRONIC TOXICITY TESTS have a PERCENT EFFECT at the IWC greater than ~~**10**~~ **25** percent.
>
> A discharge has REASONABLE POTENTIAL to cause or contribute to an excursion above the acute toxicity water quality objectives specified in Section III.B.2.b, if any of the ACUTE TOXICITY TESTS result in a "fail" at the IWC, or if any of the ACUTE TOXICITY TESTS have a PERCENT EFFECT at the IWC greater than ~~**10**~~ **25** percent."

Additionally, CVCWA disagrees with the provision in Section IV.B.2.b, which is mirrored in other sections, that those POTWs authorized to discharge at a rate equal to or greater than 5 million gallons per day (mgd) will automatically be required to have chronic toxicity effluent limits. Although a marginal improvement over the 1 mgd threshold proposed in the 2012 version of the draft Toxicity Provisions, CVCWA requests that this language be modified to state that all POTWs be allowed to perform a reasonable potential analysis to determine the need for such effluent limits, consistent with USEPA regulations contained in 40 C.F.R. § 122.44(d). If that change is not made, we request in the alternative that the Toxicity Provisions be modified to state that the Reasonable Potential assumption will apply only for the first NPDES permit renewal following adoption of the Toxicity Provisions, and that all POTWs shall be allowed to perform reasonable potential analyses to determine the need for chronic toxicity effluent limits in subsequent NPDES permit renewals.

CVCWA's preferred revisions for Page 14, Section IV. 2.b.i.

> i. Non-Storm water NPDES Dischargers Required to Conduct Reasonable
> Potential Analysis for Chronic Toxicity.

> **~~Except for POTW dischargers authorized to discharge at a rate equal to or greater than 5.0 MGD, a~~** **A**ll NON-STORM WATER NPDES DISCHARGERS shall conduct a REASONABLE POTENTIAL analysis for chronic toxicity, pursuant to the procedures specified in Section IV.B.2.b.iii, for review and approval by the PERMITTING AUTHORITY. **~~A REASONABLE POTENTIAL analysis for chronic toxicity is not required for POTW dischargers authorized to discharge at a rate equal to or greater than 5.0 MGD, because the PERMITTING AUTHORITY shall include an effluent limitation for these dischargers pursuant to Section IV.B.2.e.~~**

CVCWA's recommended edits for page 16, Section IV.c., first paragraph, are as follows:

> All NON-STORM WATER NPDES DISCHARGERS that demonstrate REASONABLE POTENTIAL for chronic toxicity **~~and all POTW dischargers that are authorized to discharge at a rate equal to or greater than 5.0 MGD~~** shall conduct monitoring for compliance with the chronic toxicity MDEL and MMEL. All NON-STORM WATER NPDES DISCHARGERS that demonstrate REASONABLE POTENTIAL for acute toxicity shall conduct monitoring for compliance with the acute toxicity MDEL and MMEL. The compliance monitoring for the MDEL and MMEL includes ROUTINE MONITORING and MMEL COMPLIANCE TESTS.

CVCWA's recommended edits for page 21, IV.B.2.e.i.A, first paragraph with a similar recommended change for the MMEL on page 22, subsection B, are as follows:

> Except when the MOST SENSITIVE SPECIES does not include the survival ENDPOINT the PERMITTING AUTHORITY shall include the following MDEL in the NPDES permit if REASONABLE POTENTIAL is demonstrated for chronic toxicity in accordance with the provisions specified in Section IV.B.2.b**~~, or if a POTW is authorized to discharge at a rate equal to or greater than 5.0 MGD~~**:

Currently, Section IV.B.2.b.iii requires five years of reference, all toxicity test data generated within five years prior to permit issuance, reissuance, renewal, or reopening (to address toxicity requirements) that is representative of effluent quality during discharge conditions shall be evaluated in determining REASONABLE POTENTIAL. The phrase "during discharge conditions" is unclear in this section, because past discharge conditions may not be representative of current discharge conditions, especially in the cases of treatment plant upgrades or additional controls within a plant. As such, CVCWA recommends the paragraph be modified in two ways, here and in other similar

provisions of the document, first by adding the word "valid" before of toxicity testing and second by clarifying that the evaluation is based on current data.

> "All **valid** toxicity test data generated within five years prior to permit issuance, reissuance, renewal, or reopening (to address toxicity requirements) that is representative of **current** effluent quality during discharge conditions shall be evaluated in determining REASONABLE POTENTIAL."

3. Test Methods and Test Endpoints

The use of the TST statistical approach proposed in the Toxicity Provisions involves the presumption that samples tested (i.e. all effluents, all ambient waters) are toxic and then relies on toxicity test results to show that they are not. This creates a presumption that depends on the toxicity testing methods being well established and consistently implemented to yield reproducible results among multiple testing laboratories. The attached Toxicity Report identifies a number of instances where significant variability in test results occurs using standard test methods and outlines suggested best practices to promote consistency in methodology. The report emphasizes the need to reduce test variability to ensure that the WET testing program is cost-effective. In that regard, changes to the proposed Toxicity Provisions are needed to add greater emphasis regarding the use of reliable and reproducible toxicity test methods.

In Section III, CVCWA requests that the proposed language addressing the "Interaction of Toxicity Provisions with Narrative and Numeric Toxicity Water Quality Objectives" be modified to establish consistent requirements regarding the Permitting Authority's discretionary capability to use alternative test organisms, test endpoints and test methods to derive effluent limitations in the application of narrative objectives.

The requested language change is as follows. In Section III.B.4 Page 4, fourth paragraph, add the following after the last sentence:

> "In exercising its discretion, the Permitting Authority shall carry the burden of demonstrating that test methods and test endpoints are reliable, repeatable, and reproducible through a process which includes, but is not limited to, documentation of test protocols, test acceptability criteria, and data quality objectives and inter-laboratory comparisons."

The importance of this was highlighted with the Southern California Coastal Research Project for the Stormwater Coalition in Southern California in 2016, which prior to the test were following non-standardized methods. However, inter-laboratory tests for blank samples for a variety of aquatic toxicity tests, including *Hyalella* and *Ceriodaphnia*, resulted in high levels of variability among the laboratories for these two species. In the second round of testing, after further consistency in method approach

were placed on the Surface Water Ambient Monitoring Program (SWAMP) protocols, the participating laboratories were able to produce much more consistent results with *Hyalella*. Additionally, testing where suppliers and timing vary still needs to be assessed to demonstrate that the method is robust enough to produce consistent results.

With regard to the test organisms and test methods specified in Table 1 of the proposed Toxicity Provisions, similar attention should be placed on the *Ceriodaphnia dubia* short term chronic reproduction test. As described in the white paper produced for CASA (Larry Walker Associates, 2018), ongoing issues persist regarding the *Ceriodaphnia dubia* reproduction test. (See CASA comments and attachments.) Those issues include variability in test results among laboratories and determination of toxicity in non-toxic samples.

Accordingly, CVCWA requests that language be added to the proposed Toxicity Provisions to limit the use of *Ceriodaphnia dubia* short term chronic reproduction tests in NPDES permits pending resolution of various testing method issues. It is also requested that the State Water Board and Regional Water Boards seek State funds to partner with the regulated community to design and implement the necessary studies to improve this test method.

4. Water Quality Objectives and the Null Hypothesis

As currently written, the proposed numeric water quality objectives are written in such a way that the State Water Board is proposing to deem all inland surface waters as toxic. CVCWA is very concerned about the consequences of such an action.

At minimum we recommend removing the paragraphs on page 2 for both acute and chronic objectives that read:

> "Attainment of the water quality objective is demonstrated by conducting CHRONIC TOXICITY TESTING as described in Section IV.B.1.b and rejecting this NULL HYPOTHESIS in accordance with the TEST OF SIGNIFICANT TOXICITY (TST) statistical approach described in Section IV.B.1.c. When the NULL HYPOTHESIS is rejected, the ALTERNATIVE HYPOTHESIS is accepted in its place, and there is no exceedance of the chronic toxicity water quality objective. Failing to reject the NULL HYPOTHESIS (referred to as a "fail") is equivalent to an exceedance of the chronic toxicity water quality objective."

Additionally, in Section III.B.2 on page 2, the wording of the proposed numeric toxicity objectives is not phrased in plain English and is difficult to understand. It is requested that the language describing the proposed objectives be modified to be more understandable to the public and regulated entities.

5. Compliance Monitoring and Sensitivity Screening

For smaller POTWs in the Central Valley, the Compliance Monitoring and Sensitivity Screening elements of the proposed Toxicity Provisions are burdensome and will add significant costs. CVCWA has several recommendations concerning these sections, which are organized in the order in which the provisions appear in the Toxicity Provisions.

- Species Sensitivity Screening & Species Use

The Toxicity Provisions require all NPDES Dischargers to conduct Species Sensitivity Screening as if they had never conducted WET testing or species sensitivity testing previously. In the Central Valley, that is not the case. Many of our POTWs have tested all three species on a regular basis for years, and have conducted screening as part of their latest permit. CVCWA recommends that the language concerning the effective date as to when screening data be considered be removed.

Page 10, Section IV.B.1.e. Reporting:

"Results obtained from **valid** toxicity tests shall be reported to the PERMITTING AUTHORITY as either a "pass" or a "fail," and the PERCENT EFFECT at the IWC for each endpoint. The results and any required supporting data shall be submitted in the format specified by the PERMITTING AUTHORITY."

Starting on page 12 of Section 2.IV.b.2.a., subsections i and ii:[2]

"All NON-STORM WATER NPDES DISCHARGERS shall conduct a SPECIES SENSITIVITY SCREENING for chronic toxicity either prior to, or within 18 months after the first issuance, reissuance, renewal, or reopening (to address toxicity requirements) of the permit ~~**after the effective date of these TOXICITY PROVISIONS**~~. The PERMITTING AUTHORITY may require a SPECIES SENSITIVITY SCREENING for chronic toxicity prior to every subsequent…"

On Page 13, Section IV.B.2.a.iii., for both chronic and acute sensitive species screening for non-continuous discharges, evaluating over a calendar year may not be the appropriate metric since discharges may be limited to a season or condition. CVCWA recommends the language for both chronic and acute screening be modified to:

For NON-CONTINUOUS DISCHARGERS, the four sets of testing shall be evenly distributed across the CALENDAR YEAR**, or during a period representative of the discharge quality,** to the extent feasible.

[2] Similar language is also used in other subsequent sections and should be modified consistently with the proposed language here.

On Page 14, Section IV.B.2.a.iv, CVCWA recommends that the last paragraph be divided to give Regional Water Boards Executive Officers flexibility during a permit term if most sensitive species cannot be used. The paragraphs would then be stand-alone.

> When the SPECIES SENSITIVITY SCREENING is conducted within 18 months of the issuance, reissuance, renewal, or reopening (to address toxicity requirements) of the permit after the effective date of these TOXICITY PROVISIONS, then the PERMITTING AUTHORITY shall specify in the NPDES permit a species as the MOST SENSITIVE SPECIES until the SPECIES SENSITIVITY SCREENING is conducted. The NPDES permit shall indicate the method of determining the MOST SENSITIVE SPECIES from the SPECIES SENSITIVITY SCREENING, and a provision indicating that the Executive Director or Executive Officer may select and document the species determined to be the MOST SENSITIVE SPECIES from the SPECIES SENSITIVITY SCREENING test.
>
> **The** PERMITTING AUTHORITY shall specify in the NPDES permit that when that species cannot be used, such as when discharger encounters unresolvable test interference or cannot secure a reliable supply of test organisms, the Executive Director or Executive Officer may specify the next applicable species as the MOST SENSITIVE SPECIES and document that determination.

- Timing of Routine Monitoring and Compliance tests

CVCWA appreciates that these Toxicity Provisions include a monthly median for chronic toxicity. However, basing a chronic toxicity limit on a single sample does not accurately reflect chronic toxicity conditions. That is based on long-term exposure of four days. Accordingly, the decision to impose a chronic toxicity effluent limitation should not be based on a single sample. Additionally, CVCWA is concerned that the Toxicity Provisions would allow the imposition of up to a half-dozen whole effluent toxicity limitations. The Toxicity Provisions do not appear to direct the Regional Water Boards to consider which limitation is most stringent, and then to apply only those limitations. The Regional Water Boards currently consider all potential WET limitations and select the most stringent, which is both protective of water quality and does not expose dischargers to unnecessary liability. CVCWA requests that the Toxicity Provisions be revised to clearly establish that Regional Water Boards use their discretion in applying only the most restrictive effluent limits in permits, rather than every potential effluent limit related to whole effluent toxicity.

CVCWA is very concerned about the requirements that the median monthly effluent limitations (MMELs) be conducted within a calendar month. As described in the BACWA's comment letter on the Toxicity Provisions, there are serious logistical issues with conducting three tests in one month. This is especially difficult when a fourth test for the following month is expected and may need to be taken adjacent to the prior

months test due to timing issues. Because most of our members do not receive final laboratory reports until three weeks after instigation of the report, our member agencies anticipate they will need to: (1) have organisms ready for two additional tests, (2) sample and transport samples to the laboratory, and (3) have the laboratory possibly start analyzing the second test at significant cost when the test ultimately may not be necessary because of the timeframe for testing associated with some of the most common species.

Furthermore, the proposed three-sample Monthly Median does not allow sufficient time associated in situations where: (1) the test acceptability criteria (TAC) is not met, (2) contract laboratory is experiencing a backlog that is outside of the discharger's control, and (3) urgent operational changes that can cause the sampling event already in progress to be aborted or re-scheduled until the treatment plants are back in normal operating condition. These scenarios may make it impossible to take the three samples within a calendar month. Because of these concerns, CVCWA recommends policy be revised to implement at least a 6-week cycle (commencement to commencement of samples) for the three-sample Median. This allows one of the compliance samples to count as the next month's sample, which saves costs. It would also avoid unnecessary costs associated with preparing, taking, and partially analyzing samples that may not be needed while still providing a reasonable method to determine compliance.

Section IV.2.c. in the first paragraph and multiple places in the section should be modified as follows:

> "The discharger shall conduct at least one CHRONIC TOXICITY TEST every CALENDAR MONTH during which there is expected to be at least 15 days of discharge. A sample for the ROUTINE MONITORING test shall be taken at a time that would allow corresponding MMEL COMPLIANCE TESTS to be initiated within **six weeks of the initiation of** ~~**the same CALENDAR MONTH as**~~ the ROUTINE MONITORING test."

- Monitoring Frequency for Toxicity and Associated Costs

The frequency of monitoring in the draft Toxicity Provisions are a significant increase over current permitting practices. This will result in substantial cost increases on dischargers. Many Central Valley POTWs with NPDES are very high-level treatment facilities. Of the 76 POTWs evaluated in CVCWA's study, over 50 of these POTWs are expected to see increases in the level of monitoring required – most increasing from quarterly or semiannually monitoring to monthly monitoring.

For example, the City of Davis, who was evaluated as part of the economic analysis for these proposed toxicity provisions (see Staff Report at Page 245, Table 9-1, showing a *decrease* in cost of $15,000), anticipate that the actual costs are expected to triple their cost for testing if the monitoring frequency under the toxicity provision is applied as compared to the City's newly adopted NPDES permit (NPDES Permit Order No. R5-2018-0085). The increase in costs listed above do not include acute toxicity testing if required by the Central Valley Water Board.

Although the draft provisions allow for reduce monitoring, the requirements for reduce monitoring under the draft Toxicity Provisions do not recognize the years of toxicity testing that POTWs have already completed and are incredibly burdensome and do not recognize plant upgrades or efforts that were taken to address toxicity if it was identified. CVCWA recommends that the level of reduction not be specified and the use of historical data be considered when determining frequency.

Other Comments

1. The definition of aquatic toxicity in Section III.B (page 1) of the proposed Provisions includes reference to "physical agents" as potential causes of adverse responses of aquatic organisms, in addition to chemical agents. This definition is atypical and may cause confusion in the implementation of the Toxicity Provisions. It is requested that the following language be used in place of the first line of the proposed definition:

"*Aquatic toxicity is the adverse effects of contaminants in aquatic ecosystems.*"

2. Section IV.B.1.b (Toxicity Test Methods): It is stated on page 7 that, while test methods listed in Table 1 specify a minimum number of replicates, additional test replicates may be conducted to increase test sensitivity and confidence in the results. This raises the question of the validity of test results performed at the minimum level of replication and whether a repeated test at increased replication should be required to confirm important results, i.e. findings of toxicity leading to potential 303(d) listings or findings triggering TRE efforts. Please address this issue in the proposed Provisions, as appropriate.

3. Section IV.B.2: It has been stated in public workshops that it is the intention of the State Water Board that the proposed Toxicity Provisions limit the establishment of acute toxicity effluent limits in the NPDES permits issued to POTWs. To firmly implement this intention, it is requested that the language of the Provisions be modified to state that the imposition of acute toxicity effluent limits should be an exception, and that Regional Water Boards (i.e. the Permitting Authority) shall be

required to provide special documentation in the NPDES permit fact sheet to justify such limits.

4. Section IV.B.2.e. Effluent Limitations. The Toxicity Provisions establish the structure of having both chronic and acute effluent limits, which deviates from the approach taken for priority pollutants in the SIP. The concern exists that a single sample result may result in multiple violations, i.e. that a test results will lead to a violation of both acute toxicity and chronic toxicity effluent limits, especially for test methods that have both mortality and sublethal endpoints. It is requested that an explanation be provided to demonstrate how this circumstance will be avoided. In the event it is found that this circumstance may occur, it is requested that changes be made to avoid episodes of multiple compliance jeopardy associated with either a single sample result or results for a single month.

5. Section IV.B.2.c. contains provisions on page d 16 -18, that the permitting authority shall specify the day of the month that corresponds to the start of the calendar month etc. Please note that other programs also use these terms and utilizing the same start date may not always be practical. Determining the appropriate start date is something that needs to be worked out between the POTW and the laboratory. We oppose the permitting authority to specify the exact dates for routine monitoring without solid justification.

6. It is unclear how the State Water Board intends to utilize the glossary in the ISWBE Plan – if as a standalone glossary to the Toxicity Provisions or combined with other terms from different programs. Our concern is the appropriateness of some of the terms that contradict with other programs, policies or new provisions. CVCWA recommends this be clear in the final policy and that should the applicability go beyond just the Toxicity Provisions, the State Water Board release a draft of the Glossary showing current terms and the newly proposed terms as it would sit in the ISWBE Plan.

We appreciate your consideration of these comments. If you have any questions, or if CVCWA can be of further assistance, please contact me at eofficer@cvcwa.org or (530) 268-1338.

Sincerely,

Debbie Webster

Debbie Webster,
Executive Officer

Enclosure

Attachment A

DECEMBER 2018

CENTRAL VALLEY CLEAN WATER ASSOCIATION

Toxicity Special Study:

Phase I Study Report

submitted by

LARRY WALKER ASSOCIATES

ROBERTSON - BRYAN, INC.
Solutions for Progress

PACIFIC ECORISK

This page is intentionally left blank.

Table of Contents

List of Appendices

Table of Tables

Table of Figures

List of Acronyms

ADWDF	Average dry weather design flow
BACWA-LC	Bay Area Clean Water Association – Laboratory Committee
BMI	Benthic macroinvertebrate
CASA	California Association of Sanitation Agencies
CETTP	Complex Effluent Toxicity Testing Program
CIWQS	California Integrated Water Quality System
CV	Coefficient of variation
CVCWA	Central Valley Clean Water Association
DCWWTP	Deer Creek Wastewater Treatment Plant
DMR-QA	Discharge Monitoring Report-Quality Assurance
DMW	Dilute mineral water
EC_{25}	Effect concentration where 25 percent of test organisms exhibit an effect
EPAMH	United States Environmental Protection Agency moderately hard
I/I	Inflow/infiltration
IC_{25}	Inhibition concentration where 25 percent of test organisms exhibit inhibition
IC_{50}	Inhibition concentration where 50 percent of test organisms exhibit inhibition
IWC	Instream waste concentration
LACSD	Sanitation Districts of Los Angeles County
LC_{50}	Lethal concentration where 50 percent of test organisms exhibit mortality
LOEC	Lowest Observed Effect Concentration
MDEL	Maximum daily effluent limitation
mgd	million gallons per day
MMEL	Maximum monthly effluent limitation
NELAP	National Environmental Laboratory Accreditation Program

NOEC	No observed effect concentration
NPDES	National Pollutant Discharge Elimination System
PMSD	Percent minimum significant difference
POTW	Publicly-owned treatment works
QA/QC	Quality assurance/quality control
SCCWRP	Southern California Coastal Water Research Project
SETAC	Society of Environmental Toxicology and Chemistry
SIP	*Policy for Implementation of Toxics Standards for Inland Surface Waters, Enclosed Bays, and Estuaries of California*
TIE	Toxicity Identification Evaluation
TRE	Toxicity Reduction Evaluation
TSD	Technical Support Document for Water Quality Based Toxics Control
TST	Test for Significant Toxicity
TU_c	Chronic toxicity units
USEPA	United States Environmental Protection Agency
WPCF	Water Pollution Control Facility
WERF	Water Environment Research Foundation
WET	Whole effluent toxicity
YCT	Yeast Cerophyll®-trout

This page is intentionally left blank.

Executive Summary

1.0 INTRODUCTION

Publicly-owned treatment works (POTWs) are required, through the National Pollutant Discharge Elimination System (NPDES) Program and permits, to conduct periodic chronic whole effluent toxicity (WET) testing to determine if treated effluent impacts aquatic life and the ecosystem in the receiving water of the effluent discharge. If a chronic toxicity test indicates toxicity in the effluent discharge (typically by exceeding a chronic toxicity trigger), a POTW must follow specific procedures, outlined by the United States Environmental Protection Agency (USEPA) and state water boards, for verifying the observed effluent toxicity, investigating and identifying the cause(s) of the effluent toxicity, and implementing the appropriate control measure(s) to mitigate/eliminate the toxicity in the effluent discharge.

Low-level effects appear to represent a significant fraction of chronic toxicity trigger exceedances based on available chronic bioassay data for POTWs in the Central Valley. As a starting point based on Central Valley 2018 NPDES permit language for chronic toxicity testing, the working definition used for low-level chronic toxicity is chronic bioassay test results with a chronic toxicity trigger of ≤2 chronic toxicity units (TU_c) *and* a percent reduction of less than 25 percent when comparing the receiving water concentration (instream waste concentration [IWC]) sample (e.g., typically 100 percent effluent where the chronic toxicity trigger is 1 TU_c or >1 TU_c) with the control water (e.g., laboratory water, receiving water) sample. Because of the follow-up investigation required in NPDES permits, these exceedances result in significant expenditures for accelerated testing and, in many cases, subsequent Toxicity Reduction Evaluation (TRE), which can include costly Toxicity Identification Evaluation (TIE) studies. Initial information received from POTWs indicates that accelerated testing, TRE studies, and/or TIE testing may not result in the identification of the cause(s) of observed low-level effluent chronic toxicity.

The following issues were identified by POTWs regarding the nature of observed low-level effects, the ability to determine the cause of observed low-level effects, and the impact that these observed effluent effects may have on the beneficial uses of the receiving water:

- Flexibility in test conditions that may be used by a laboratory (as allowed by USEPA test guidelines) and/or natural variability in the sensitivity of test organisms that exists between different laboratory culture or batches of organisms;
- Variability of test results for the same effluent sample among different laboratories that may lead to different conclusions as to whether an effluent indicates toxicity or not;
- Sensitivity of conventional TIE testing that may not be able to identify the causative toxicant, or even toxicant class, due to a low and/or non-persistent toxicity signal; and
- Uncertainty that there is a measureable effect to aquatic life species or beneficial uses in the receiving water from the discharge as a result of the low-level effects occasionally being observed in effluent bioassay tests.

Unless an effluent is truly toxic and the POTW is able to identify and mitigate the cause of observed low-level chronic effects, the POTW may not be able to exit the TRE process in a

conventional manner and thus may be required to continue with a TRE for an extended time period, which requires significant resources.

The purpose of this study is to better focus POTW and Central Valley Regional Water Quality Control Board (Central Valley Water Board) efforts and resources on the reasonable protection of beneficial uses in the receiving water through the examination of additional scientific and regulatory responses to low-level effects observed in effluent chronic bioassay tests. The goals of this study are to:

- Determine the frequency with which chronic toxicity test exceedances are observed by Central Valley POTWs conducting chronic three-species bioassay testing and whether these exceedances may be classified as low-level effects;
- Evaluate the efficacy of TREs and TIEs in resolving indications of effluent toxicity;
- Document potential variability in sub-lethal test endpoints;
- Identify, if possible, the level of sub-lethal effects in chronic WET tests that correlate to measurable effects to aquatic life in the receiving water; and
- Develop a preliminary conceptual model that identifies the factors that may result in indications of toxicity during chronic toxicity testing and factors that are anticipated to increase the likelihood that chronic toxicity test results will correlate with observable effects in the receiving water.

This study is funded by a special project group of the Central Valley Clean Water Association (CVCWA), which represents POTWs in the Central Valley. CVCWA's mission is to represent the interests of wastewater agencies in the Central Valley in regulatory matters that balance the need for environmental protection based on sound scientific information with a fair and reasonable economic basis. As of August 2018, a total of 22 Central Valley POTWs have contributed to this special project.

This executive summary is organized into the following sections with the corresponding sections in the Phase I Study Report (in parentheses) that provide additional detail:

- Section 1: Introduction (**Section 1**)
- Section 2: Characterization of Chronic Toxicity Test Results from Central Valley POTWs (**Section 4**)
- Section 3: Variability in Sub-lethal Endpoints (**Section 5**)
- Section 4: Relationship Between Toxicity Testing and Aquatic Ecosystem Impacts (**Section 6**)
- Section 5: Draft Conceptual Model (**Section 7**)
- Section 6: Summary of Key Findings and Recommendations (**Section 8**)

2.0 CHARACTERIZATION OF CHRONIC TOXICITY TEST RESULTS FROM CENTRAL VALLEY POTWS

There are approximately 77 POTWs located in the Central Valley that are regulated by the Central Valley Water Board under the NPDES Program. These POTWs provide sewerage services for over 7 million people, manage and treat wastewater generated by domestic, commercial, industrial, and other sources, and range from dischargers with an average dry weather design flow (ADWDF) of 0.026 million gallons per day (mgd) to 181 mgd. Effluent treated by these POTWs is discharged into various types of receiving waters, including agricultural conveyances, creeks, rivers, streams, and, in limited circumstances, lakes. Additionally, treated effluent is being increasingly utilized as recycled water to supplement and augment water supplies.

2.1 POTW and Data Sources Background

For this study, chronic toxicity test reports (i.e., routine chronic toxicity testing, accelerated testing, TRE/TIE testing) were obtained for 66 Central Valley POTWs for the period of January 2011 to March 2017 through the State Water Resources Control Board's (State Water Board) California Integrated Water Quality System (CIWQS) database, communications with Central Valley Water Board staff, and directly from POTWs. The purpose of obtaining full chronic toxicity test reports instead of only compiling TU_c results, which are readily available in CIWQS, was to document the level of "effect" observed in the effluent relative to the control (e.g., *C. dubia* reproductive inhibition, *S. capricornutum* cell growth inhibition). Ideally, this will provide information to allow future evaluation of effects levels at which follow-up TRE efforts were able to successfully identify the cause(s) of indicated toxicity.

2.2 Routine Chronic Toxicity Testing Characterization for Central Valley POTWs

As discussed above, the Central Valley Water Board requires chronic toxicity testing as one element of its program to meet the requirements of the Basin Plan (and Clean Water Act) in protecting aquatic life beneficial uses of surface waterbodies. In Central Valley NPDES permits (generally found in Attachment E [Monitoring and Reporting Program]), POTWs are required to conduct periodic chronic toxicity testing for three freshwater species: fathead minnow (*Pimephales promelas*), water flea (*Ceriodaphnia dubia*), and green alga (*Selenastrum capricornutum*, also known as *Raphidocelis subcapitata*). *P. promelas* is tested for 7 days for the growth and survival endpoints, *C. dubia* is tested for 6-8 days for survival and reproduction, and *S. capricornutum* is tested for 96 hours for growth. The testing frequency varies depending on the size of the POTW and ranges from once per NPDES permit term (e.g., once every five years) to monthly. Most Central Valley POTWs are required to conduct chronic toxicity testing either on an annual or quarterly basis during periods of discharge to surface receiving waters. All sample collection and testing must adhere to USEPA's *Short-term Methods for Estimating Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms, Fourth Edition* (EPA/821-R-02-013, October 2002).

Test results are currently required to be reported in terms of chronic toxicity units (TU_c). As of March 2017 (the end of the chronic toxicity data set for Central Valley POTWs evaluated for this

study), Central Valley POTWs were required to calculate TU_c as 100 divided by the "No Observable Effect Concentration" (NOEC). The NOEC is statistically determined as the lowest effluent test concentration in which there was no observed significant effect.

Test results are compared to a chronic toxicity trigger in the NPDES permit to determine compliance with the toxicity objectives. Most Central Valley POTWs chronic toxicity triggers do not consider dilution and therefore have a numeric toxicity monitoring trigger of 1 TU_c on the basis of the NOEC as determined by hypothesis testing. The Central Valley Water Board has the discretion to provide a higher numeric toxicity monitoring trigger if receiving water dilution is considered. As a result, several Central Valley POTWs have higher numeric toxicity monitoring triggers (e.g., 4 TU_c, 8 TU_c). Exceedance of a chronic toxicity trigger requires follow-up accelerated testing.

While the chronic toxicity testing requirements discussed above are the current requirements, the State Water Board is currently in the process of developing a Statewide Toxicity Policy to establish water quality objectives for aquatic toxicity and a statistical approach for assessing toxicity in POTW and other effluents and receiving waters. The proposed policy would be included in the statewide *Water Quality Control Plan for Inland Surface Waters, Enclosed Bays, and Estuaries* and would supersede current requirements in Basin Plans. The proposed policy would establish numeric water quality objectives for both acute and chronic toxicity and establish a program of implementation to control toxicity. Under the most recent publicly-available proposal[1], attainment of the water quality objective would be demonstrated by rejecting the null hypothesis, which states that "the ambient receiving water is toxic because the test organism adverse response in ambient receiving water sample is significantly different from the test organism response in the control water sample." The statistical analysis will be conducted using the Test of Significant Toxicity (TST) (USEPA 2010).

Under the proposed Statewide Toxicity Policy, MDELs and MMELs for chronic toxicity are being proposed with thresholds applicable to the difference between IWC and control toxicity tests with sub-lethal endpoints (e.g., 50 percent effect; no statistically significant effect) and represents a shift from the historic focus on TU_c metrics in chronic toxicity test results. This shift results in the use of the "effect" metric (percent difference from control) in determining compliance with the numeric water quality objectives.

As of September 2018, the State Water Board has not yet re-released a public review draft of the Statewide Toxicity Policy.

2.2.1 Routine Chronic Toxicity Testing Characterization

Nearly 1,000 routine chronic toxicity tests (i.e., testing conducted according to NPDES permit monitoring frequencies) were obtained for each of the three test species (e.g., *P. promelas*, *C.*

[1] California State Water Resources Control Board. *Summary of Proposed Toxicity Provisions* (April 2017). https://www.waterboards.ca.gov/water_issues/programs/state_implementation_policy/tx_ass_cntrl.shtml, Last accessed February 7, 2018.

dubia, and *S. capricornutum*) and evaluated in this analysis. In distinguishing routine chronic toxicity test data and other chronic toxicity test data (e.g., accelerated testing, TRE testing), routine chronic toxicity tests were generally identified by when the sampling and toxicity test occurred in relation to the required NPDES permit monitoring frequency for each POTW. There were instances where routine chronic toxicity testing overlapped with accelerated testing and/or TRE/TIE testing. For these situations, the tests (including baseline tests for TRE/TIE testing) were classified as routine chronic toxicity tests for the purpose of this analysis. (Chronic toxicity tests identified as accelerated testing, TRE/TIE testing, or ones conducted at a frequency greater than the required NPDES permit monitoring frequency were excluded from the baseline characterization.) In reviewing the routine chronic toxicity testing frequency requirements, it is estimated that the data set utilized in this analysis is missing approximately 200 routine chronic toxicity tests based on the expected NPDES permit testing frequency for the POTWs included in this study. This means that approximately 80 to 85 percent of the expected total routine chronic toxicity tests between January 2011 and March 2017 were included in the data set for this analysis.

It is critical to note that some Central Valley POTWs have chronic numeric toxicity triggers that considered receiving water dilution, which resulted in increasing triggers from 1 TU_c up to a maximum of 16 TU_c. Of the 66 Central Valley POTWs included in this study, 7 POTWs have a chronic toxicity trigger higher than 1 TU_c. With the exception of the City of Woodland, which receives a chronic toxicity trigger of 2 TU_c for *S. capricornutum* only, the remaining 58 Central Valley POTWs (59 Central Valley POTWs including the City of Woodland for *P. promelas* and *C. dubia*) evaluated in this study have a chronic toxicity trigger of 1 TU_c.

2.2.1.1 Chronic Toxicity Test Results Based on TU_c Metric

An overall summary of chronic toxicity trigger exceedances for Central Valley POTWs during routine chronic toxicity testing for each test endpoint is presented in **Table ES-1** along with the number of POTWs impacted by chronic toxicity trigger exceedances.

Table ES-1. Central Valley POTWs Chronic Toxicity Trigger Exceedances, January 2011 to March 2017.

Test Organism/Endpoint	Total Number of Chronic Toxicity Tests	Number of Chronic Toxicity Trigger Exceedances (%)	Number and Percent of Central Valley POTWs Impacted
Central Valley POTWs with Chronic Toxicity Trigger = 1 TU_c (1)			
Pimephales promelas (survival)	832	4 (0.5%)	2 (3.4%)
Pimephales promelas (growth)	834	20 (2.4%)	15 (25.4%)
Ceriodaphnia dubia (survival)	818	9 (1.1%)	7 (11.8%)
Ceriodaphnia dubia (reproduction)	820	131 (16.0%)	38 (64.4%)
Selenastrum capricornutum (growth)	835	76 (9.1%)	29 (49.2%)
Central Valley POTWs with Chronic Toxicity Trigger > 1 TU_c (2)			
Pimephales promelas (survival)	128	2 (1.6%)	1 (14.3%)
Pimephales promelas (growth)	128	2 (1.6%)	1 (14.3%)
Ceriodaphnia dubia (survival)	137	0 (0.0%)	0 (0.0%)
Ceriodaphnia dubia (reproduction)	138	17 (12.3%)	3 (42.9%)
Selenastrum capricornutum (growth)	152	2 (1.3%)	1 (12.5%)

(1) There are 59 Central Valley POTWs that have a chronic toxicity trigger of 1 TU_c for all test organisms. Prior to December 1, 2014, the City of Woodland had a chronic toxicity trigger of 1 TU_c for *S. capricornutum*. Chronic toxicity data collected for the City of Woodland prior to this date were included in this subset of data.

(2) There are 7 Central Valley POTWs that have a chronic toxicity trigger of greater than 1 TU_c for all test organisms. After December 1, 2014, the City of Woodland had a chronic toxicity trigger of 2 TU_c for *S. capricornutum*. Chronic toxicity data collected for the City of Woodland after this date were included in this subset of data.

As stated previously, NPDES permits adopted by the Central Valley Water Board from the end of 2014 through 2016 only required chronic toxicity testing of 100 percent effluent and the control water(s). As a result, chronic toxicity test results were often reported as greater than 1 TU_c when chronic toxicity was observed in the effluent sample. This revised approach for testing resulted in data depicted as greater than 1 TU_c, which confounds attempts to define or determine whether the results were "low-level" or not. The levels of reported chronic toxicity trigger exceedance for *C. dubia* reproduction and *S. capricornutum* growth and exceedance frequencies are presented in **Figures ES-1** and **ES-2**, respectively, for the available data for the entire study period.

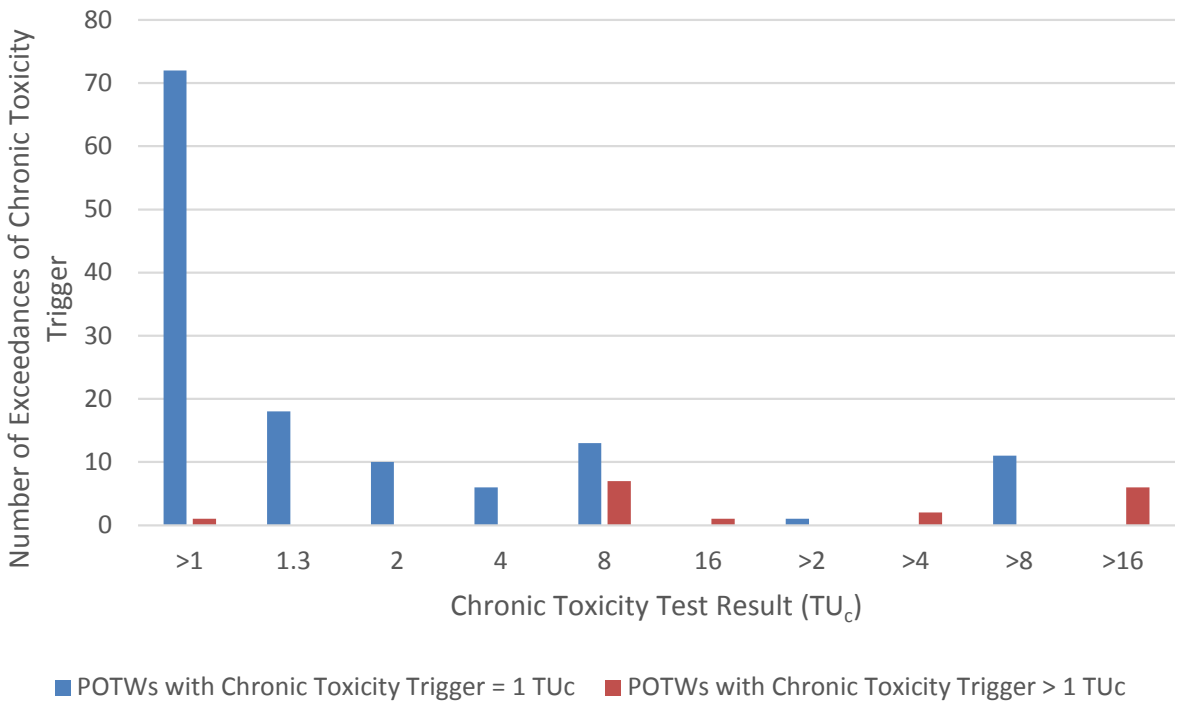

Figure ES-1. Chronic Toxicity Trigger Exceedances for *Ceriodaphnia dubia* Reproduction for Central Valley POTWs, January 2011-March 2017.

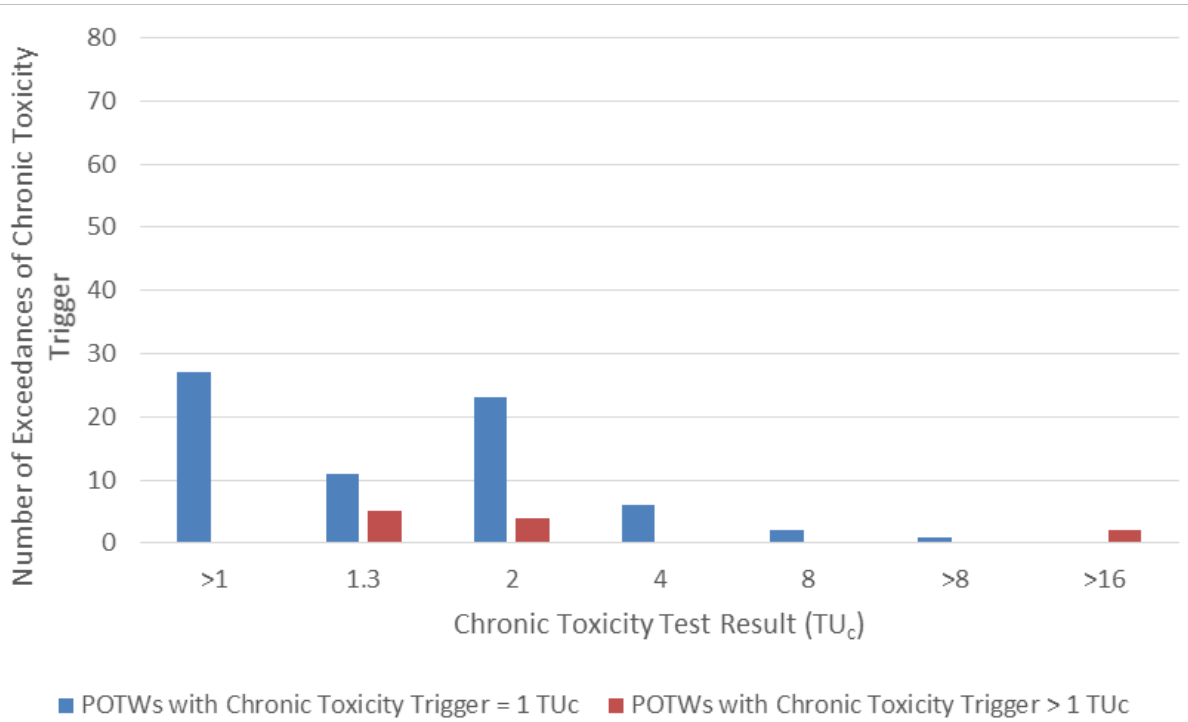

Figure ES-2. Chronic Toxicity Trigger Exceedances for *Selenastrum capricornutum* Growth for Central Valley POTWs, January 2011-March 2017.

Key Findings

The following key findings were observed in the baseline analyses evaluating exceedances of the chronic toxicity trigger:

- Most chronic toxicity trigger exceedances for Central Valley POTWs were observed for *C. dubia* reproduction and *S. capricornutum* growth for POTWs that have a chronic toxicity trigger of 1 TU_c. Approximately two-thirds of the POTWs experienced at least one chronic toxicity trigger exceedance for *C. dubia* reproduction and half of the POTWs observed at least one exceedance for *S. capricornutum* growth during the time period assessed for this study. Most of the remaining chronic toxicity trigger exceedances for *C. dubia* reproduction and *S. capricornutum* growth were 1.3 or 2 TU_c, which means that, in those cases, toxicity was generally only observed in the 100 percent effluent sample, and was not present when the sample was diluted by 25 or 50 percent, respectively.
- Chronic toxicity was only sporadically observed for *P. promelas* survival and growth and *C. dubia* survival. The remaining analyses did not focus on these test species endpoints.
- For Central Valley POTWs that have a chronic toxicity trigger of greater than 1 TU_c, the *C. dubia* reproduction test endpoint was exceeded more frequently than the other chronic toxicity endpoints, and affected approximately half of the POTWs in this category. Central Valley POTWs that have dilution (and higher chronic toxicity triggers) had significantly fewer exceedances compared to POTWs that have a chronic toxicity trigger of 1 TU_c.

2.2.1.2 Chronic Toxicity Test Results Based on Percent Effect Metric

Although the TU_c metric has been used to evaluate chronic toxicity of effluent, the State Water Board's 2012 proposed Statewide Toxicity Plan proposed to no longer use the TU_c metric as the benchmark for evaluating chronic toxicity data and assessing compliance with the Basin Plan water quality objectives. The proposed Statewide Toxicity Plan recommends use of a percent "effect" metric to assess the magnitude of chronic toxicity by comparing differences in toxicity test results between the effluent and control water(s). Chronic toxicity test results from Central Valley POTWs were also characterized based on the percent effect where there was also an exceedance of the chronic toxicity trigger to determine if there was any correlation between chronic toxicity trigger exceedance and the observed percent effect. A summary of the percent effect on the subset of chronic toxicity data where an exceedance of the chronic toxicity trigger was observed for *C. dubia* reproduction and *S. capricornutum* is presented in **Tables ES-2** and **ES-3**, respectively.

Table ES-2. Percent Effect for *Ceriodaphnia dubia* Reproduction Chronic Toxicity Trigger Exceedances for Central Valley POTWs, January 2011 to March 2017.

Control Water	Number of Chronic Toxicity Trigger Exceedances	<25% Reduction	25-50% Reduction	50-75% Reduction	>75% Reduction
Central Valley POTWs with Chronic Toxicity Trigger = 1 TU_c (1)					
Laboratory water (2)	104	31 (29.8%)	48 (46.2%)	13 (12.5%)	11 (10.6%)
Receiving water	27	6 (22.2%)	14 (51.9%)	4 (14.8%)	3 (11.1%)
Central Valley POTWs with Chronic Toxicity Trigger > 1 TU_c (3)					
Laboratory water	4	1 (25.0%)	2 (50.0%)	1 (25.0%)	0 (0.0%)
Receiving water	13	0 (0.0%)	0 (0.0%)	0 (0.0%)	13 (100%)

(1) There are 59 Central Valley POTWs that have a chronic toxicity trigger of 1 TU_c for *C. dubia* reproduction.
(2) There was one toxicity test report that indicated laboratory water was used, but the raw reproduction data were not included, which results in the columns not summing to the total.
(3) There are 7 Central Valley POTWs have a chronic toxicity trigger greater than 1 TU_c for *C. dubia* reproduction.

Table ES-3. Percent Effect for *Selenastrum capricornutum* Growth Chronic Toxicity Trigger Exceedances for Central Valley POTWs, January 2011 to March 2017.

Control Water	Number of Chronic Toxicity Trigger Exceedances	<25% Reduction	25-50% Reduction	50-75% Reduction	>75% Reduction
Central Valley POTWs with Chronic Toxicity Trigger = 1 TU_c (1)					
Laboratory water (2)	38	20 (52.6%)	9 (23.7%)	5 (13.2%)	4 (10.5%)
Receiving water	38	13 (34.2%)	11 (28.9%)	11 (28.9%)	3 (7.9%)
Central Valley POTWs with Chronic Toxicity Trigger > 1 TU_c (3)					
Laboratory water	0	–	–	–	–
Receiving water	2	2 (100%)	0 (0.0%)	0 (0.0%)	0 (0.0%)

(1) There are 59 Central Valley POTWs that have a chronic toxicity trigger of 1 TU_c for *S. capricornutum*. Before December 1, 2014, the City of Woodland had a chronic toxicity trigger of 1 TU_c for *S. capricornutum*. Chronic toxicity data collected for the City of Woodland prior to this date were included in this subset of data.
(2) There are 7 Central Valley POTWs that have a chronic toxicity trigger of greater than 1 TU_c for *S. capricornutum*. After December 1, 2014, the City of Woodland has a chronic toxicity trigger of 2 TU_c for *S. capricornutum*. Chronic toxicity data collected for the City of Woodland after this date were included in this subset of data.

Key Findings

The following key findings were observed in the baseline analyses evaluating the percent effect between the effluent sample and the control:

- When chronic toxicity was indicated for *C. dubia* reproduction (i.e., the chronic toxicity trigger was exceeded), approximately a quarter of those tests showed effects less than 25 percent, and approximately three-quarters of those tests showed effects less than 50 percent.
- When laboratory water was used as the control for *S. capricornutum* testing, approximately half of the tests exceeding the chronic toxicity trigger had effects less than 25 percent, with three-quarters of those exceeding the trigger showing effects less than 50 percent. When the receiving water was used as the control, approximately one-third of the tests with chronic toxicity trigger exceedances showed effects less than 25 percent, while approximately two-thirds showed effects less than 50 percent.
- It is generally acknowledged that a persistent effect of at least 30 to 50 percent in a sample when compared to the control is desirable for successful TIE implementation (USEPA 1996, USEPA 2007). The results for *C. dubia* where more than one-quarter of the chronic toxicity trigger exceedances observed during the study period would not be anticipated to yield successful TIE outcomes. For *S. capricornutum*, more than one-third of the chronic toxicity trigger exceedances fall in this category. These levels of toxicity may have resulted in POTWs expending extensive resources to investigate low-level toxicity that would not be anticipated to be resolved.

2.2.1.3 Other Factors Assessed

In addition to the baseline chronic toxicity test assessment discussed above, trend analyses were conducted on the following factors to determine if they had an impact on chronic toxicity results for *C. dubia* reproduction and *S. capricornutum* growth:

- Control water;
- Temporal impacts;
- Seasonality;
- POTW treatment level;
- Nitrogen treatment; and
- Disinfection methodology.

Key Findings

The following key findings were observed in these analyses:

- Exceedances of the chronic toxicity trigger for *S. capricornutum* growth for these POTWs occurred equally between receiving water and laboratory water controls for POTWs with a chronic toxicity trigger of 1 TU_c. One concern that has been raised is that the use of receiving water for *S. capricornutum* growth may have biostimulatory effects (e.g., higher concentrations of nutrients or other characteristics that may allow for the test organisms to thrive in this control water). This test variability may result in indications of toxicity in the effluent as it is compared to the control that promotes higher growth.
- In a year over year analysis of chronic toxicity trigger exceedances for *C. dubia* reproduction, there appeared to be an increasing number of exceedances as a percentage of the number of chronic toxicity test performed as well as in the number of POTWs impacted for Central Valley POTWs that have a chronic toxicity trigger of 1 TU_c.

- In Central Valley POTWs that have a chronic toxicity trigger of 1 TU_c, increased nutrient treatment appeared to reduce the frequency in which exceedances of the chronic toxicity trigger occurred for *S. capricornutum* growth.
- In evaluating *S. capricornutum* growth-related toxicity trigger exceedances based on POTW disinfection methodology, the chronic toxicity trigger was exceeded at twice the rate for ultraviolet light disinfection when compared to chlorination, and approximately three-quarters of the POTWs using ultraviolet light disinfection experienced an exceedance of the chronic toxicity trigger compared to approximately one-fifth of the POTWs using chlorination disinfection. This indicates that the type of disinfection methodology may negatively impact *S. capricornutum* reproduction.

Other Findings

- An analysis was also conducted to evaluate if seasonality impacted chronic toxicity trigger exceedances. The results of that analysis found that chronic toxicity trigger exceedances did not vary from season to season in the full chronic toxicity testing data set. However, this does not preclude seasonal impacts on individual POTWs.
- While higher levels of wastewater treatment also appear to result in fewer chronic toxicity trigger exceedances for *C. dubia* reproduction (Chi-square at $\alpha = 0.5$, p-value = 0.010) for POTWs with a chronic toxicity trigger of 1 TU_c, approximately three-quarters of the exceedances are attributed to the City of Davis, which operated a land-based secondary treatment system and underwent treatment plant upgrades during the period of the data set. Excluding the City's data set from this analysis resulted in POTWs with only secondary treatment having a chronic toxicity trigger exceedance rate of approximately 10.2 percent.
- In general, the level of treatment provided by a POTW did not appear to impact the frequency of observed chronic toxicity trigger exceedances for either *C. dubia* reproduction or *S. capricornutum* growth when the data set was adjusted to exclude chronic toxicity results from the City of Davis. The City of Davis operated a land-based treatment system, which was unique compared to other secondary-level treatment POTWs, and underwent a wastewater treatment plant upgrade during the data period. The results from the City of Davis significantly skewed the data set for other POTWs operating secondary level treatment.
- There did not appear to be a significant difference in the type of disinfection methodology impacting *C. dubia* reproduction when the City of Davis data set is excluded. It is important to note that nearly every POTW using ultraviolet light disinfection has observed an exceedance of its chronic toxicity trigger for *C. dubia* compared to about half of those POTWs using chlorination.

2.2.1.4 Key Data Set Issues

The following key issues and items of note regarding the routine chronic toxicity test data set that was used in this evaluation were identified:

- Chronic toxicity test reports were not consistently uploaded to CIWQS prior to about 2011 or 2012. Subsequent to this period, until March 2017 (the end date for the data set

used in this analysis), most routine chronic toxicity test reports conducted by larger POTWs were uploaded to CIWQS and were available for this analysis.

- Smaller POTWs were not required or were on a different time schedule to submit electronic data to CIWQS. As such, chronic toxicity test reports and data for smaller Central Valley POTWs were not readily available.
- Through the time period of the data set, permitting approaches to chronic toxicity evaluations were changed to address information learned from chronic toxicity testing findings and effectiveness and impacts of the testing and TREs, including costs associated with conducting follow-up investigation of toxicity trigger exceedances. This made it difficult to evaluate certain trends associated with chronic toxicity trigger exceedances using TU_c. Examples of the changes in permitting approaches include the following:
    - The control water requirements for chronic toxicity testing differed among POTW NPDES permits. Older NPDES permits typically specified the control water to which effluent results were compared. More recent NPDES permits allow the POTW to use either laboratory or receiving water for the control water.
    - The requirement for a dilution series for chronic toxicity testing changed during the data period evaluated. In NPDES permits adopted prior to the end of 2014 and from 2017 forward, POTWs were required to conduct routine chronic toxicity tests with a dilution series using either laboratory or receiving water. In NPDES permits adopted from the end of 2014 to the end of 2016, POTWs that have a numeric toxicity monitoring trigger of 1 TU_c were typically not required to conduct chronic toxicity testing with a dilution series unless the POTW was undergoing a TRE.

2.3 Accelerated Testing Characterization for Central Valley POTWs

Accelerated testing is required if a toxicity test indicates an exceedance of the chronic toxicity trigger, and typically consists of four chronic toxicity tests conducted once every two weeks. If none of those four accelerated chronic toxicity tests exceed the chronic toxicity trigger, the POTW can cease accelerated testing and resume routing testing in accordance with the testing frequency required in its NPDES permit. However, if any of the four accelerated tests indicates an exceedance of the chronic toxicity trigger, the POTW is required to initiate a TRE. NPDES permits allow dischargers who readily identify the cause of toxicity during accelerated testing to take corrective actions and reinitiate accelerated testing to verify that there is no longer an indication of toxicity in the effluent.

Beginning with NPDES permits adopted in December 2017, which is outside the data period evaluated as part of this study, the Central Valley Water Board made modifications to its approach for requiring accelerated testing. POTWs were required to conduct accelerated testing only if both the chronic toxicity trigger was exceeded and the results in the 100 percent effluent were at least 25 percent less than the control. This approach is expected to help resolve some of the resource-intensive follow-up issues with potential low-level chronic toxicity that are discussed in this report.

2.3.1 Accelerated Testing Characterization

Based on the characterization of routine chronic toxicity testing, there were approximately 263 instances (i.e., chronic toxicity trigger exceedances) where POTWs may have needed to initiate accelerated testing to follow-up on an indication of toxicity. From the data review and compilation, there were approximately 362 chronic toxicity tests that were characterized as accelerated testing and utilized for this analysis. Of the accelerated testing reports that were evaluated as part of this study, 21 were associated with *P. promelas*, 282 were associated with *C. dubia*, and 80 were associated with *S. capricornutum*. It should be noted that the total accelerated test reports do not sum to the total accelerated testing reports reviewed as part of this study because there were instances where a POTW conducted accelerated testing on multiple test organisms at the same time.

From the available accelerated testing data, a breakdown of the type of follow-up conducted by the Central Valley POTWs is presented in **Figure ES-3**. Approximately 40 percent of the POTWs proceeded with a TRE because there was a second exceedance of the chronic toxicity trigger during accelerated testing. Fourteen percent of the POTWs did not experience a second chronic toxicity trigger exceedance during accelerated testing and returned to routine chronic toxicity testing while three percent were still in the process of conducting accelerated testing at the end of the data period evaluated.

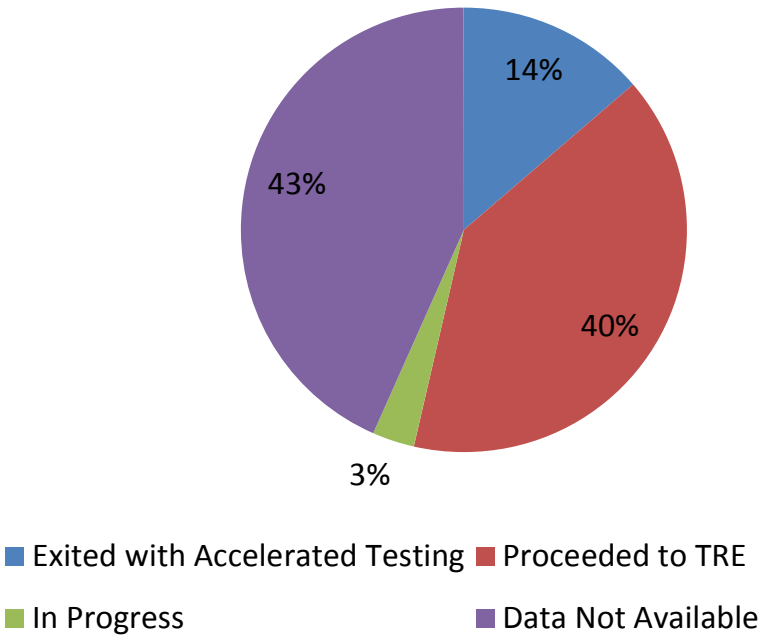

Figure ES-3. Breakdown of Follow-up Activities to Chronic Toxicity Trigger Exceedances

For 43 percent of the chronic toxicity trigger exceedances, accelerated testing results were not available for evaluation in this study. Some reasons for this may include any or all of the following:

- One of the challenges in reviewing accelerated testing data was that POTWs used different nomenclature for how toxicity tests subsequent to routine monitoring were labeled. In some situations, accelerated tests were labeled as TRE- or TIE-related testing or vice versa. In the data collection phase of this study, some POTWs provided follow-up information in cover letters of required reporting to the Central Valley Water Board without toxicity test attachments. Because accelerated testing is a follow-up effort outside typically routine and information that is required to be uploaded to CIWQS, the test reports may have been directly submitted to the POTW's case handler at the Central Valley Water Board;
- Where possible, some POTWs overlapped accelerated testing with routine monitoring. This approach typically helps save costs associated with the extra monitoring, although it introduced some complexity to this data analysis.
- There may have been an easily discernible cause of the toxicity that was identified and corrected;
- The discharge may have been seasonal or discontinued and monitoring was not appropriate after the discharge ceased; and/or
- A routine chronic toxicity test sample may have been contaminated or identified as not being representative of the discharge.

To help further refine the analysis of the outcomes of accelerated testing, an additional study would need to focus on this portion of unknown data to better understand how POTWs handled accelerated testing (e.g., how many returned to routine chronic toxicity testing, how many entered the TRE process).

2.4 Toxicity Reduction Evaluation/Toxicity Identification Characterization for Central Valley POTWs

As noted above, a TRE is typically initiated when effluent toxicity is observed to be persistent, which is most often demonstrated when repeated chronic toxicity tests indicate the presence of toxicity during accelerated testing. The goal of a TRE is to reduce or eliminate effluent toxicity. Specifically, a TRE is:

> *"A site-specific study conducted in a step-wise process designed to identify the causative agent(s) of effluent toxicity, isolate the sources of toxicity, evaluate the effectiveness of toxicity control options, and then confirm the reduction in effluent toxicity."* (USEPA 1991)

In *Toxicity Reduction Evaluation Guidance for Municipal Wastewater Treatment Plants* (USEPA 1999), USEPA presents an acceptable protocol for implementing a TRE. Typically, this is the protocol that is most often followed or referenced by NPDES permittees and permit writers when a TRE is initiated (see **Figure 4**). The TRE protocol is generally divided into six component parts, or tiers, including:

1. Information and Data Acquisition (Tier 1)
2. Facility Performance Evaluation (Tier 2)
3. Toxicity Identification Evaluation (Tier 3)
4. Toxicity Source Evaluation (Tier 4)

5. Toxicity Control Evaluation (Tier 5)
6. Toxicity Control Implementation (Tier 6)

It should be noted that while the protocol is presented in a linear fashion (i.e., sequentially numbered tiers), a TRE can be successfully concluded at any stage of the process, and some tiers can be omitted, conducted concurrently, or implemented out of sequence. Moreover, the TRE protocol is not a mandate, and activities other than those described in the protocol that lead to effluent toxicity resolution may also be considered. As clarified by USEPA in *Clarifications Regarding Toxicity Reduction and Identification Evaluations in the National Pollutant Discharge Elimination System Program* (USEPA 2001):

> "*Any activities that result in consistently reducing toxicity to an acceptable level may be considered TRE activities.*"

2.4.1 Toxicity Reduction Evaluation Characterization

Unlike obtaining self-monitoring report data from a central repository such as CIWQS, obtaining detailed information on TRE experiences, TRE strategies, and TRE outcomes was challenging, as this information is not submitted in a consistent manner or stored in a centralized location. Information for 39 completed TREs were compiled based on contributions from the special study consultant team and Central Valley Water Board and through a data solicitation directed to CVCWA member agencies. These 39 completed TREs represented the experience of 25 different POTWs; 9 POTWs provided information on more than one TRE.

Four of the 39 TREs were concluded after it was determined that the results of the toxicity testing falsely indicated toxicity. In all four cases, there was a notable test interference or test method protocol deviation identified. Test interferences included a) plating of algae cells (i.e., sticking of algae cells to the walls of the test container) in the *S. capricornutum* test, b) pathogen effects on *P. promelas*, c) biostimulatory receiving water used in *C. dubia* tests, and d) test method protocol deviation for dissolved oxygen in *C. dubia* tests. In each case, evidence of the interference of method deviation was used as the basis to conclude each TRE. Because these four TREs were not related to actual effluent toxicity, they were excluded from the data set that was evaluated.

The test organisms and endpoints subject to the 35 TREs are presented in **Figure ES-4**. Four of the 35 TREs were associated with POTWs that have a chronic toxicity trigger of greater than 1 TU_c.

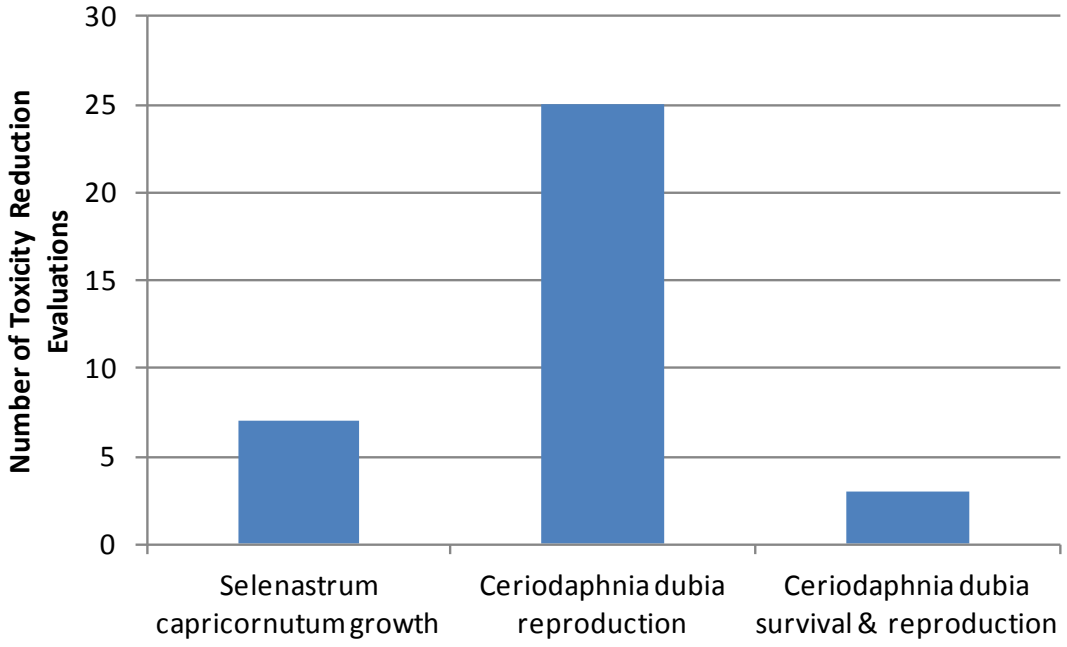

Figure ES-4. Central Valley POTW Toxicity Reduction Evaluation Test Organism and Endpoint.

Generalizations as to the identified cause of toxicity can be made, in addition to generalizations as to which step of the TRE process was most informative in advancing the TRE's conclusion. Information was also obtained regarding the duration of the TREs. This evaluation is limited to these generalizations, as each TRE is unique in regard to the POTW affected, the circumstances triggering the TRE, and the level of experience of the parties managing the TRE investigation.

For the 35 completed TREs compiled, identified causes of toxicity were categorized as follows (**Figure ES-5**):

- POTW maintenance or construction;
- Collection system maintenance or construction;
- Significant industrial user (SIU or other non-domestic or non-commercial user);
- POTW chemical usage;
- An identified contaminant; or
- Unknown cause.

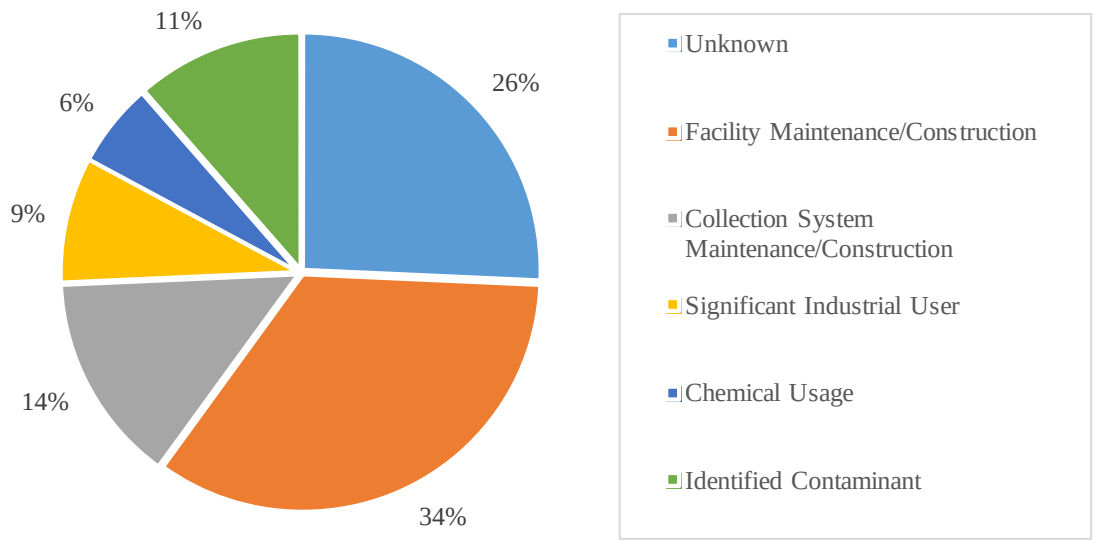

Figure ES-5. Identified Cause of Toxicity in 35 Toxicity Reduction Evaluations for Central Valley Publicly-Owned Treatment Works.

The TRE step that was most informative in advancing the TRE's conclusion is depicted in **Figure ES-6**.

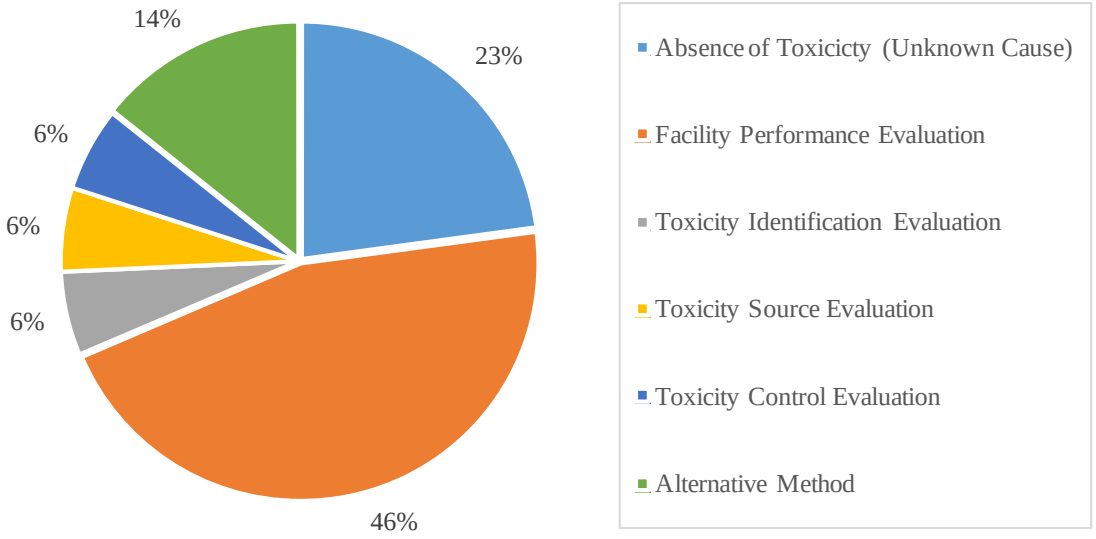

Figure ES-6. Tier Most Informative as to Successful Conclusion of 35 Toxicity Reduction Evaluations.

As shown in **Figure ES-7**, of the 35 TREs evaluated, 77 percent were concluded within 1 year of initiation, with 9 percent of TREs exceeding 3 years. Of the three TREs that exceeded three years, two were concluded after substantial infrastructure improvements were constructed, including installation of new treatment processes. In these two cases, the duration of the TRE is partially explained by time involved in the design, financing, and/or construction of new or upgraded treatment facilities.

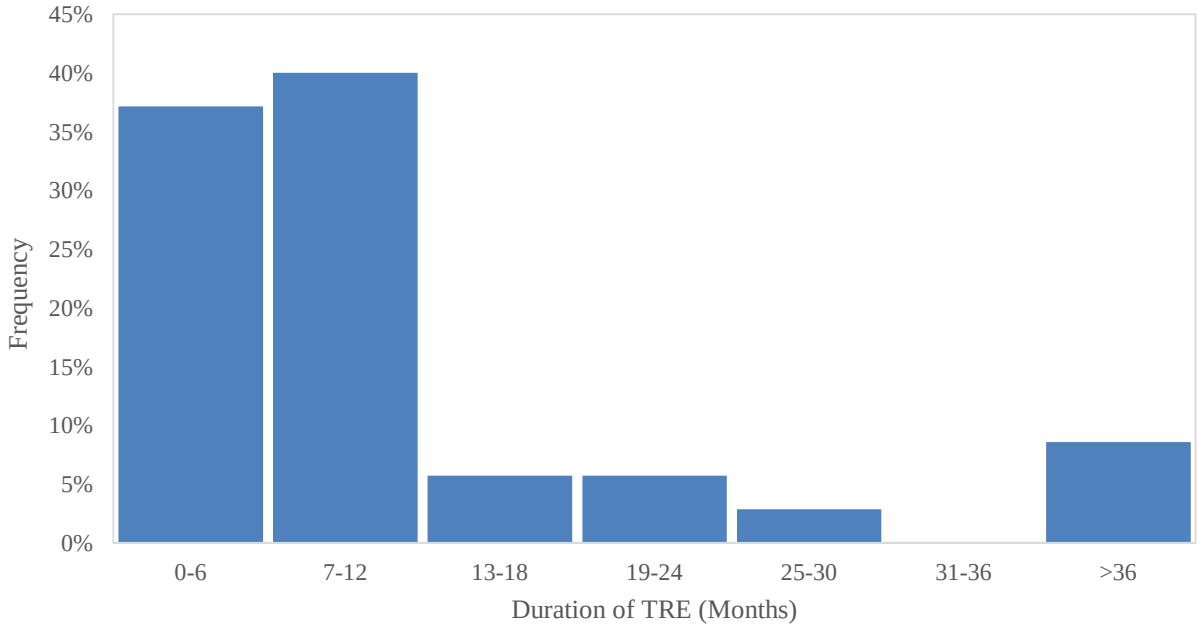

Figure ES-7. Duration of Toxicity Reduction Evaluations.

Key Findings

The following key findings were observed in analyzing completed TREs:

- Activities related to POTW or collection system maintenance or construction were identified as the cause of effluent toxicity in 48 percent of the TREs (i.e., 17 of 35). In addition to these cause and POTW chemical usage, about half of the chronic toxicity trigger exceedances may have been potentially controllable.
- In 26 percent of the TREs (i.e., 9 of 35), no cause for the observed toxicity could be identified. In each of those cases, toxicity was shown to be absent in multiple follow-up chronic toxicity tests. Of the nine TREs where no cause for observed toxicity could be identified, five of these TREs were associated with a relative effect (i.e., difference between the effluent sample and control) of less than 25 percent at the POTWs' IWC in at least one of the two tests required to trigger a TRE. The remaining four of these TREs were associated with a relative effect of 50 percent or greater at the POTWs' IWC. Additional information would be required to further evaluate these individual cases of an "unknown cause" in order to understand the individual drivers behind these inconclusive TREs.

- The Facility Performance Evaluation step was most influential, and in all but a single case, the accompanying TRE was concluded in less than one year.
- In contrast, the TIE step was found to be most influential in only 6 percent of TREs despite the TIE step being a component of 34 percent of all TREs. In other words, of 12 TREs that implemented the TIE step, in only 2 cases did the TIE protocol advance the TRE to a conclusion. Where TIE laboratory reports were available in the compiled CIWQS database, unsuccessful TIEs (i.e., those leading to little or no substantive information) were primarily associated with diminished or absent baseline toxicity. Most TIEs are performed on samples first demonstrated to be toxic, and thus the TIE is performed on an aged sample (typically more than seven days from sample collection). For labile contaminants subject to degradation or transformation, toxicity can diminish quickly in effluent held in storage, thus rendering the majority of TIE treatments inconclusive, as acknowledged by USEPA (2007):

  "If the baseline test does not show consistent, measurable toxicity, then one cannot perform a TIE, as the effect of the manipulations on toxicity cannot be assessed."

  While the available TIE laboratory reports were insufficient to determine the threshold of effect at which TIEs are generally successful, it is generally acknowledged by USEPA that a persistent effect of approximately 30 to 50 percent in a sample is desirable for successful TIE implementation (USEPA 1996, USEPA 2007). TIE success can also depend on the experience of the laboratory performing the TIE manipulations and the breadth of manipulations employed.

- In 23 percent of the TREs, the cause of toxicity could not be identified because effluent toxicity was no longer observed.
- In five TREs, the alternative methods such as instream bioassessment, comparative split laboratory testing, and a special request to the Central Valley Water Board to transition to a Toxicity Evaluation Study were used to conclude the TRE.

2.5 Split-Laboratory Testing and Variability Summary

One avenue that is sometimes explored by POTWs involved in a TRE is "split-laboratory" testing to determine if multiple laboratories concur on the presence or absence of observed toxicity. An analysis was conducted to assess available "split-laboratory" data to determine if this approach provided additional information into potential factors that may impact results from chronic toxicity testing. Overall, "split-laboratory" studies resulted in a moderate to high degree of agreement. For the "split-laboratory" comparisons that were performed, the greatest agreement between laboratories occurred for the *P. promelas* test. The laboratories always agreed for this test, but it is important to note that the sample size (n=4) was quite small for this protocol. The *C. dubia* test had the next highest agreement between (and among) laboratories (73-83%), and the lowest agreement between laboratories occurred with the *S. capricornutum* test (65-77%). Typically, there was a slightly greater agreement between laboratories using a comparison of the IC_{25} as when compared to the NOEC.

When only two laboratories were used in "split-laboratory" testing and the laboratories both generate acceptable test data, but are in disagreement, it is unclear which laboratory should be

used to assess compliance. A principal limitation of pair-wise "split-laboratory" testing is the lack of a consistent and defensible means of placing higher value on one test results over another when the test results are not in agreement. Development of a consistent and defensible approach to "split-laboratory" testing, if "split-laboratory" testing is to be employed with rigor, is recommended. Warren Hicks et al., (2000) recommendations should be considered in which they indicated that comparability among laboratories should be addressed by using ***several*** laboratories (emphasis added) before making a determination on whether an effluent is in compliance.

3.0 VARIABILITY IN SUB-LETHAL ENDPOINTS

There are three sources of method variability: intra-test, intra-laboratory, and inter-laboratory. Intra-test variability is the variability of the test organism response within a single test. Intra-laboratory variability is the variability of tests conducted over time within the same laboratory, which is affected by intra-test variability. Inter-laboratory variability is the variability among laboratories, which is measured by evaluating the results of different laboratories testing the same sample(s) using the same test method. Inter-laboratory variability is affected by intra-test variability and intra-laboratory variability. The factors that affect these three forms of test variability are provided in **Table ES-4**. Changes in any of the factors listed in **Table ES-4** can affect variability (e.g., such as changes in culture practices).

Table ES-4. Sources of Variability for Aquatic Toxicity Testing Methods.

Category of Variability	Sources of Variability
Intra-test variability	- Replicates: number of replicates and number of organisms per replicate - Culture quality: genetic variability, culture condition - Microbial interferences*:* epibionts (e.g., peritrichs, bacteria.)
Intra-laboratory variability	- Test conditions: selection and variability in food, control/dilution water quality/consistency, consistency of test conditions - Organism condition: culture quality (affected by conditions above) - Laboratory experience: testing facility, quality assurance/quality control (QA/QC) program, analyst training program, variability among analysts - Analyst experience: training support, adherence to protocol
Inter-laboratory variability	- Intra-test variability factors: see above - Intra-laboratory factors: see above - Differences allowed in method: source and type of food, control/dilution water, organism culture condition

Although all laboratories in California are required to be accredited by the State Water Board's Environmental Laboratory Accreditation Program, the chronic toxicity test method provides flexibility on various factors ranging from test organism culturing (e.g., food, genetic variability) to microbial interferences to laboratory analyst training. These factors, among others, can impact the conclusions developed from chronic toxicity testing and the resultant findings on determining compliance with water quality standards.

Of the three chronic test species required to be tested by Central Valley POTWs for which inter-laboratory studies have been performed, the chronic *P. pimephales* growth exhibited the highest precision, followed by the *S. capricornutum* growth endpoint; the *C. dubia* reproduction endpoint exhibited the lowest precision even though there have been more inter-laboratory studies performed for this species and endpoint over a 30-year period of time.

4.0 RELATIONSHIP BETWEEN TOXICITY TESTING AND AQUATIC ECOSYSTEM IMPACTS

While WET testing is often used to assess final effluent from POTWs, it only provides a snapshot of the effluent and characterizes it as toxic or non-toxic isolated from its actual impact on the receiving waters and aquatic life. As stated previously, the Basin Plan includes a narrative objective for toxicity. The Basin Plan continues:

> "*Compliance with this objective will be determined by analyses of indicator organisms, species diversity, population density, growth anomalies, and biotoxicity tests of appropriate duration or other methods as specified by the Regional Water Board.*"

If POTWs are unable to resolve the cause of chronic toxicity in effluent, other in-stream evaluations (e.g., indicator organisms, species diversity, population density) may need to be considered to assess if the effluent is causing toxicity in the receiving water and assess compliance with the narrative toxicity objective. A literature review was conducted to summarize and discuss the general history of studies that have examined the link between toxicity tests and aquatic ecosystem effects, with a specific emphasis on levels of sub-lethal endpoint toxicity and characteristics that do and do not correlate to measurable effects to aquatic life. It was particularly challenging to discern from literature the level(s) of sub-lethal endpoint toxicity that do or do not correlate with receiving water effects) for many reasons. Some of these reasons are related to study design, including, but not limited to, the following examples:

- Study designs vary widely;
- Not all characteristics of toxicity or bioassessment are reported in all studies;
- Statistical methods may lump lethal and sub-lethal endpoints which confounds interpretation of results; and
- Each physical site is unique and is usually affected by multiple factors beyond effluent quality.

A number of common themes were identified from the literature as follows:

1. In general, it does not appear that WET test results are reliable predictors of effects or lack of effects in the receiving water environment (Chapman 2000, Diamond 2000, Diamond et al. 2008). Specifically, intermittent and low-level toxicity, as measured by sub-lethal endpoints, does not appear to be a reliable predictor of receiving water impairments.
2. In general, ambient water toxicity testing better represents biological condition of the water body than effluent testing, and higher magnitudes of ambient toxicity are better correlated with biological effects (de Vlaming and Norberg-King 1999, Dickson 1992).

3. WET testing is better representative of instream biological effects when dilution is considered and when higher magnitudes of toxicity are present after considering dilution (de Vlaming and Norberg-King 1999, Diamond et al. 2000, City of Woodland).
4. Higher frequencies and magnitudes of WET toxicity can be generally better correlated with biological effects in a water body (Dickson 1992, de Vlaming and Norberg-King 1999, Diamond et al. 2000).
5. There is no consensus on which WET test species provide the best predictions of biological condition in the receiving water. Different studies have reached different conclusions on this matter (Diamond and Daley 2000, Diamond et al. 2008).
6. Because biological responses measured in WET tests are considered less reliable near test detection limits (de Vlaming and Norberg-King 1999), predictions of biological effects in a waterbody based on WET testing will be improved when laboratory performance and data quality for freshwater chronic WET tests is evaluated with measurement quality objectives that include the use of split test evaluation, blind positive control testing, blind negative control testing, and reference toxicants (Diamond et al. 2008).

A single, specific "level" of sub-lethal endpoint toxicity (i.e., magnitude of effect in WET tests) that correlates well to measurable effects in aquatic life could not be identified, largely because of point #1 above (in general, sub-lethal WET tests results are not reliable predictors of effects or lack of effects in the receiving environment). Although it does not appear as if a single, well-defined threshold level can be identified, criteria that are associated with improved correlation between WET test results and measurable effects on aquatic life can be identified. For example, as stated above, Diamond and Daley (2000) found that WET results for *P. promelas* correlated with instream impairments when:

1. The effluent comprised 80 percent or more of stream flow under design conditions;
2. Instream habitat quality was characterized as fair to good;
3. At least three WET tests had been conducted; and
4. A test failure rate (i.e., toxicity had been detected) of at least 25 percent had occurred.

5.0 DRAFT CONCEPTUAL MODEL

A fundamental principal of toxicology is the association of increased effect, such as impairment of reproduction, with increased concentration, or dose, of toxicant. Typically, it assumed that a causal relationship exists between the concentration of a contaminant and a measured response in the organisms. The classic concentration response would be considered a sigmoidal shaped curve, in which the response in the organism increases as the contaminant concentration increases, with more severe effects (e.g., acute survival) typically occurring at the higher concentrations and less severe responses (e.g., growth and reproduction) occurring at lower concentrations.

A conceptual model designed for this study includes all potential drivers for toxicity (low-level or higher level), with many of the drivers in the conceptual model having been identified in TRE studies performed for Central Valley POTWs. The conceptual model is divided into three elements: POTW drivers, testing laboratory drivers, and the environmental drivers of toxicity. Examples of these drivers for toxicity is presented below:

- POTW drivers of toxicity
    - Upstream sources of wastewater including industrial users, collection system maintenance and operations activities, severe infiltration and inflow, and source water;
    - POTW operations, maintenance and construction;
    - Sample contamination;
    - Background receiving water issues, including stimulatory responses from presence of additional nutrients or food and pathogens;
- Testing laboratory drivers of toxicity
    - Laboratory expertise and experience;
    - Test organism quality;
    - Test interferences;
    - Test design;
- Environmental drivers of toxicity
    - Instream conditions;
    - Laboratory variability; and
    - Observed toxicity testing results.

Based on these drivers of toxicity, a draft conceptual model is presented in **Figure ES-8**.

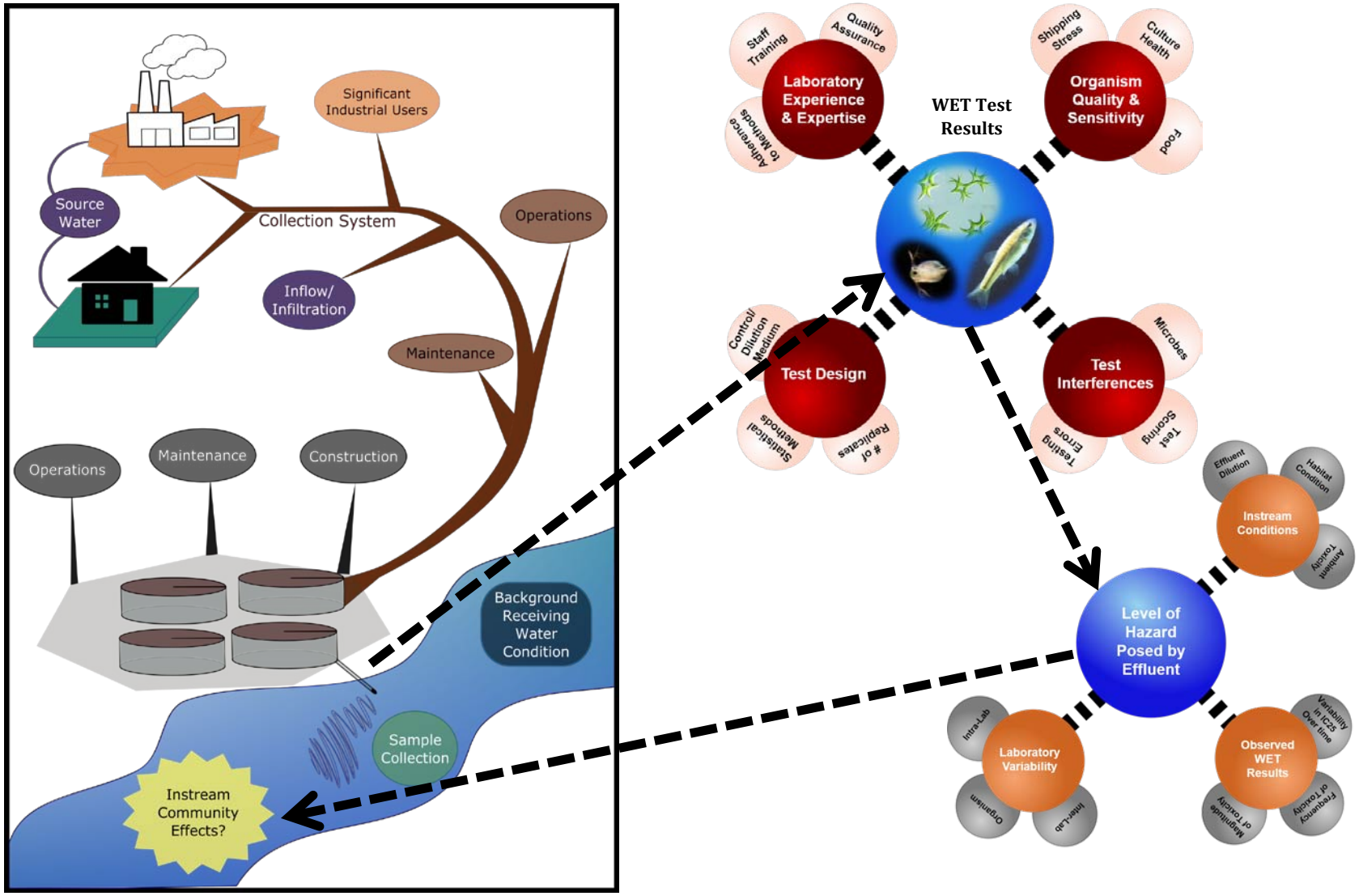

Figure ES-8. Draft Conceptual Model for Assessing Factors Influencing Chronic Toxicity Test Results and Level of Hazard Posed by Effluent to Instream Aquatic Communities.

6.0 SUMMARY OF KEY FINDINGS AND RECOMMENDATIONS

Central Valley POTWs are required to conduct periodic three-species (*P. promelas, C. dubia*, and *S. capricornutum*) chronic toxicity testing or to use the most sensitive of the three species in chronic toxicity testing to assess the impact that treated effluent may potentially have on the receiving waters and its beneficial uses, including aquatic life. To better understand the nature of the potential issues that surround exceedances of the chronic toxicity trigger, CVCWA conducted this study to characterize the extent to which low-level effects in chronic bioassay tests occur for Central Valley POTWs, identify how exceedances of the chronic toxicity trigger are resolved using the available tools developed and approved by USEPA, evaluate the efficacy of these tools, and develop a conceptual model to better understand numerous variables that can impact the outcome of a chronic toxicity test and the relationship with impairment to instream ecology. From this study, CVCWA may identify, develop, and evaluate additional tools that could be used to better resolve incidents of low-level chronic toxicity in a more effective and cost-efficient way. These additional tools would be proposed to the Central Valley Water Board to supplement existing tools that are currently used by POTWs to investigate and resolve current incidents of identifiable chronic toxicity.

6.1 Study Key Findings

The key findings of this study are discussed below.

6.1.1 Central Valley POTW Chronic Toxicity Characterization

- Central Valley POTWs with a chronic toxicity trigger of 1 TU_c
    - In reviewing chronic toxicity test data from January 2011 to March 2017, Central Valley POTWs with a chronic toxicity trigger of 1 TU_c primarily have exceedances of the chronic toxicity trigger for *C. dubia* reproduction and *S. capricornutum* growth, which are both sub-lethal endpoints. Central Valley POTWs only have isolated incidences of chronic toxicity for *P. promelas* survival and growth and *C. dubia* survival.
    - The majority of the exceedances for Central Valley POTWs with a chronic toxicity trigger of 1 TU_c was 1.3 or 2 TU_c depending on the dilution series that was utilized in the chronic toxicity test. This means that toxicity was observed only in the 100 percent effluent, but not observed in subsequent dilutions during toxicity testing.
    - POTWs using ultraviolet light disinfection observe toxicity for *S. capricornutum* growth twice as frequently as POTWs using chlorination disinfection. Nearly all POTWs utilizing ultraviolet light disinfection have experienced an exceedance of the chronic toxicity trigger for *S. capricornutum* during the data period evaluated.
    - A temporal analysis indicates that the total number of exceedances of the toxicity trigger for *C. dubia* reproduction has increased on a year to year basis.
- Central Valley POTWs with a chronic toxicity trigger greater than 1 TU_c
    - In reviewing chronic toxicity test data from January 2011 to March 2017, Central Valley POTWs with a chronic toxicity trigger greater than 1 TU_c primarily have exceedances of the chronic toxicity trigger for *C. dubia* reproduction. These

Central Valley POTWs have isolated incidences of chronic toxicity for *P. promelas* survival and growth, *C. dubia* survival, and *S. capricornutum* growth.

- Accelerated testing analysis
  - Based on the available data set, forty percent of accelerated testing conducted by POTWs lead to a TRE.
  - Based on the available data set, fourteen percent of accelerated testing conducted by POTWs did not indicate further chronic toxicity, which allowed POTWs to return to routine monitoring requirements.
  - Because of limitations to the accelerated testing data set, follow-up study of additional accelerated testing can be conducted to improve the data set and refine the understanding of the outcomes of accelerated testing.

6.1.2 Evaluate the efficacy of TREs and TIEs in resolving indications of effluent toxicity

- Toxicity Reduction Evaluation Analysis
  - For the majority of TREs that were reviewed, the TREs were resolved through a facility and/or operations review. Accelerated testing for the majority of these incidents were triggered by a relative percent difference of less than 50 percent.
  - Nearly one-quarter of the studies were eventually concluded without identifying the cause or likely cause of the toxicity. In these cases, the likely reason as to why these studies were not resolved is due to lack of persistence in toxicity.
  - TIE testing was conducted as part of 12 TREs, but in only 2 cases were TIE testing effective in identifying the cause of toxicity.
- 'Split-Laboratory' Analysis
  - 'Split-laboratory' studies resulted in a moderate to high degree of agreement with a chronic toxicity triggers. For the 'split-laboratory' comparisons that were performed, the greatest agreement between laboratories occurred for the *P. promelas* test; the laboratories always agreed for this test, but it is important to note that the sample size (n=4) was quite small for this protocol. The *C. dubia* test had the next highest agreement between (and among) laboratories (73.3 to 82.7 percent), and the lowest agreement between laboratories occurred with the *S. capricornutum* test (65 to 77 percent).
  - Typically, there was slightly greater agreement in determining compliance with the trigger between laboratories using a comparison of the IC_{25} as when compared to the NOEC.

6.1.3 Variability of Sub-Lethal Endpoints

- There are a variety of sources of variability, including numerous sources of intra-test, intra-laboratory, and inter-laboratory variability. All sources of test variability may play a role that can result in different test outcomes between/among laboratories.
- *P. promelas* growth exhibited the highest precision, followed by the *S. capricornutum* growth endpoint; the *C. dubia* reproduction endpoint exhibited the lowest precision even though there have been more inter-laboratory studies performed over an estimated 30-year period.

- There are a number of POTW, WET testing laboratory, and environmental drivers that can influence the outcome of toxicity tests. When possible, control of the POTW drivers can improve the outcome of the toxicity tests. Similarly, laboratories that have experienced technicians can reduce the influence of some drivers in the laboratory (e.g., organism quality, test interferences, and test design) as can the type of control water selected, thereby minimizing factors that confound the outcome of toxicity tests.
- Although the literature provides general sources of intra- and inter-laboratory variability, it would be exceedingly challenging to identify the specific causes of intra- and inter-laboratory variability for the 'split-laboratory' testing evaluated in this study as the compiled 'split-laboratory' testing was not designed to investigate the cause of different test outcomes.

6.1.4 Relationship Between Toxicity Testing and Aquatic Ecosystem Impacts

- A literature review performed for this study indicated that it does not appear that WET test results are reliable predictors of effects or lack of effects in the receiving water environment (Chapman 2000, Diamond 2000, Diamond et al. 2008). Specifically, intermittent and low-level toxicity, as measured by sub-lethal endpoints, does not appear to be a reliable predictor of receiving water impairments. Some studies have shown a qualitative correlation between *P. promelas* and *C. dubia* ambient toxicity tests and instream biological condition, but there is considerable debate as to whether these studies are representative of effluents and their receiving waters in general. Since it is particularly challenging to identify levels of sub-lethal endpoint toxicity that do or not correlate with receiving water effects, only general conclusions can be made.
- Ambient water toxicity testing better represents biological condition of the water body than effluent testing, and higher magnitudes of ambient toxicity are better correlated with biological effects (de Vlaming and Norberg-King 1999, Dickson 1992).
- WET testing is better representative of instream biological effects when dilution is considered (de Vlaming and Norberg-King 1999, Diamond et al. 2000, City of Woodland).
- Higher frequencies and magnitude of WET toxicity are generally better correlated with biological effects in a water body (Diamond et al. 2000; Diamond et al. 2008).
- There is no consensus on which WET test species provide the best predictions of biological condition in the receiving water. Different studies have reached different conclusions on this matter (Diamond and Daley 2000, Diamond et al. 2008).
- Because biological responses measured in WET tests are considered less reliable near test detection limits (de Vlaming and Norberg-King 1999), predictions of biological effects in a water body based on WET testing will be improved when laboratory performance and data quality for freshwater chronic WET tests is evaluated with measurement quality objectives that include the use of 'split-laboratory' test evaluation, blind positive control testing, blind negative control testing, and reference toxicants (Diamond et al. 2008).

6.2 STUDY RECOMMENDATIONS

Based on the key findings of this study summarized above, the following recommendations are made to CVCWA for Phase II or subsequent phases of the Toxicity Special Study:

- Conduct Phase II of the Toxicity Special Study, which will further evaluate whether low-level toxicity equates to adverse effects to receiving water aquatic life and beneficial uses, identify and evaluate additional tools that could be utilized by POTWs to investigate low-level indications of toxicity, finalize the conceptual model, and use the technical information compiled and evaluated to refine, expand, and strengthen the toxicity testing process applied to POTWs through NPDES permits. These additional tools can provide an alternative for POTWs to resolve low-level toxicity using methods other than TIE testing if the situation warrants such an approach.
- Refine the accelerated testing data set (e.g., conduct follow-up investigation of the 43 percent of routine chronic toxicity tests that resulted in an indication of toxicity, but toxicity test data during accelerated testing were not available for this study) to better understand the frequency in which Central Valley POTWs conduct TREs or return to routine chronic toxicity monitoring after completion of accelerated testing without a second exceedance of the chronic toxicity trigger. Follow-up investigation can be conducted for the 43 percent of routine chronic toxicity
- Recommend POTWs consider using a third laboratory when 'split-laboratory' testing results in different conclusions to resolve the toxicity and determine compliance through a weight-of-evidence approach.
- Recommend further study into potential causes and correlation of increased chronic toxicity observed for *S. capricornutum* for POTWs using ultraviolet light disinfection in comparison to chlorination-based disinfection.
- Encourage the formation of the SETAC Issues Group and the continuation of SCCWRP-led studies related to addressing and improving inter-laboratory precision for *C. dubia* testing.
- Recommend that POTWs currently conducting a TRE, determine whether the receiving water is currently being impacted by low-level toxicity. In determining whether this would be a useful or beneficial exercise, the criteria at the end of **Section 6** can be reviewed to determine the likelihood that test results will correlate with impacts in the receiving water. It can be reasonably assumed, though not guaranteed, from the literature review conducted that the more criteria that are satisfied, the greater the hazard posed by the effluent to instream biological condition and the more likely that WET results will correlate with receiving water biological condition.

Section 1. Introduction

Publicly-owned treatment works (POTWs) are required, through the National Pollutant Discharge Elimination System (NPDES) Program and permits, to conduct periodic chronic whole effluent toxicity (WET) testing to determine if treated effluent may be negatively impacting aquatic life and the ecosystem in the receiving water of the effluent discharge. If a chronic toxicity test indicates toxicity in the effluent discharge (typically by exceeding a chronic toxicity trigger), a POTW must follow specific procedures, outlined by the United States Environmental Protection Agency (USEPA) and state water boards, for verifying the observed effluent toxicity, investigating and identifying the cause(s) of the effluent toxicity, and implementing the appropriate control measure(s) to mitigate/eliminate the toxicity in the effluent discharge.

The California State Water Resources Control Board (State Water Board) is currently in the process of developing a Statewide Toxicity Policy to establish water quality objectives for aquatic toxicity and a statistical approach for assessing toxicity in POTW effluent and receiving waters. The proposed policy would be included in the statewide *Water Quality Control Plan for Inland Surface Waters, Enclosed Bays, and Estuaries* and would supersede current requirements in Basin Plans.

Based on available chronic bioassay data for POTWs in the Central Valley, low-level effects appear to represent a significant fraction of chronic toxicity trigger exceedances. These exceedances result in significant expenditures for accelerated testing and, in many cases, subsequent Toxicity Reduction Evaluation (TRE), which can include costly Toxicity Identification Evaluation (TIE) studies, as presently required in NPDES permits. Anecdotal information from POTWs indicates that accelerated testing, TRE studies, and/or TIE testing may not result in the identification of the cause(s) of observed low-level effluent chronic toxicity.

The issues to be addressed in this study are: (1) whether expenditures of time and resources using existing tools to address low-level chronic toxicity represent an effective allocation of limited public resources to provide reasonable protection of beneficial uses of receiving waters; and (2) whether new approaches and tools are needed to address low-level effects that are observed in routine bioassay tests.

As stated in **Section 1.2**, a goal of this study is to define the definition of low-level chronic toxicity in order to improve the success rate of TRE/TIE studies in identifying the cause(s) of indicated toxicity. As a starting point based on Central Valley 2018 NPDES permit language for chronic toxicity testing, the working definition used for low-level chronic toxicity is chronic bioassay test results with a chronic toxicity trigger of ≤ 2 chronic toxicity units (TU_c) *and* a percent reduction of less than 25 percent when comparing the receiving water concentration (instream waste concentration [IWC]) sample (e.g., typically 100 percent effluent where the chronic toxicity trigger is 1 TU_c or >1 TU_c) with the control water (e.g., laboratory water, receiving water) sample.

1.1 STUDY BACKGROUND

The highly sensitive nature of sub-lethal endpoints in chronic toxicity tests enables the identification of low-level differences in test organism response between effluent and control water. Under the current requirements and approved chronic toxicity testing procedures, when statistically significant (at a nominal error rate [alpha] of 0.05 when using hypothesis testing), these differences are deemed to indicate "toxicity". The response required, according to NPDES permits, for such results when the indication is in the effluent is initiation of accelerated testing and, potentially, the initiation of a TRE if accelerated testing also indicates a continued unacceptable level of effluent "toxicity". The TRE is a multi-phased approach for identifying the cause(s) of toxicity that starts with relatively simple background evaluations (e.g., facility performance evaluation and review of existing effluent quality data). If a cause for the toxicity is not resolved through the simple background investigations, the TRE proceeds to more complex efforts that can include TIEs, parallel treatment and testing methods that attempt to identify the constituent class (e.g., organics, metals) causing toxicity and may identify the specific compound(s) causing toxicity. TREs may be concluded after implementing control measures at the POTW or within its service area to mitigate the toxicity.

The following are fairly common concerns that have been voiced by POTWs regarding the nature of observed low-level effects, the ability to determine the cause of observed low-level effects, and the impact that these observed effluent effects may have on the beneficial uses of the receiving water:

- Flexibility in test conditions that may be used by a laboratory (as allowed by USEPA test guidelines) and/or natural variability in the sensitivity of test organisms that exists between different laboratory culture or batches of organisms;
- Variability of test results for the same effluent sample among different laboratories that may lead to different conclusions as to whether an effluent indicates toxicity or not;
- Sensitivity of conventional TIE testing that may not be able to identify the causative toxicant, or even toxicant class, due to a low and/or non-persistent toxicity signal; and
- Uncertainty that there is a measureable effect to aquatic life species or beneficial uses in the receiving water from the discharge as a result of the low-level effects occasionally being observed in effluent bioassay tests.

Unless an effluent is truly toxic and the POTW is able to identify and mitigate the cause of observed low-level chronic effects, the POTW may not be able to exit the TRE process in a conventional manner and thus may be required to continue with a TRE for an extended time period, which requires significant resources.

1.2 PURPOSE AND GOALS OF PHASE I STUDY

The purpose of this study is to better focus POTW and Central Valley Regional Water Quality Control Board (Central Valley Water Board) efforts and resources on the reasonable protection of beneficial uses in the receiving water through the examination of additional scientific and regulatory responses to low-level effects observed in effluent chronic bioassay tests. The goals of this Phase I Study are to:

- Determine the frequency with which chronic toxicity test exceedances are observed by Central Valley POTWs conducting chronic three-species bioassay testing and whether these exceedances may be classified as low-level effects;
- Evaluate the efficacy of TREs and TIEs in resolving indications of effluent toxicity;
- Document the potential variability in chronic toxicity testing for sub-lethal test endpoints;
- Identify, if possible, the level of sub-lethal effects in chronic WET tests that correlate to measurable effects to aquatic life in the receiving water; and
- Develop a preliminary conceptual model that identifies the factors that may result in indications of toxicity during chronic toxicity testing and factors that are anticipated to increase the likelihood that chronic toxicity test results will correlate with observable effects in the receiving water.

A potential Phase II Study to follow-up this study would be focused on further evaluating whether low-level toxicity equates to adverse effects to receiving water aquatic life and beneficial uses, identifying and evaluating additional tools that could be utilized by POTWs to investigate low-level indications of toxicity, finalizing the conceptual model, and using the technical information compiled and evaluated to refine, expand, and strengthen the toxicity testing process applied to POTWs through NPDES permits, and possibly the State Water Board's Statewide Toxicity Policy.

This study is funded by a special project group of the Central Valley Clean Water Association (CVCWA), which represents POTWs in the Central Valley. CVCWA's mission is to represent the interests of wastewater agencies in the Central Valley in regulatory matters that balance the need for environmental protection based on sound scientific information with a fair and reasonable economic basis. As of August 2018, a total of 22 Central Valley POTWs have contributed to this special project.

1.3 PHASE I STUDY REPORT OUTLINE/CONTENTS

This study report is organized into the following sections:

- Section 1: Introduction
- Section 2: Central Valley Publicly-Owned Treatment Works Background
- Section 3: Chronic Toxicity Testing Requirements
- Section 4: Characterization of Chronic Toxicity Test Results from Central Valley POTWs
- Section 5: Variability in Sub-lethal Endpoints
- Section 6: Relationship Between Toxicity Testing and Aquatic Ecosystem Impacts
- Section 7: Draft Conceptual Model
- Section 8: Summary of Key Findings and Recommendations
- Section 9: References

This page is intentionally left blank.

Section 2. Central Valley Publicly-Owned Treatment Works Background

There are approximately 77 POTWs located in the Central Valley that are regulated by the Central Valley Water Board under the NPDES Program. These POTWs provide sewerage services for over 7 million people, manage and treat wastewater generated by domestic, commercial, industrial, and other sources, and range from dischargers with an average dry weather design flow (ADWDF) of 0.026 million gallons per day (mgd) to 181 mgd. Effluent treated by these POTWs is discharged into various types of receiving waters, including agricultural conveyances, creeks, rivers, streams, and, in limited circumstances, lakes. Additionally, treated effluent is being increasingly utilized as recycled water to supplement and augment water supplies.

2.1 PUBLICLY-OWNED TREATMENT WORKS TREATMENT LEVELS

Central Valley POTWs can be generally classified into one of three treatment categories (secondary, advanced secondary, or tertiary).[2] Secondary treatment of wastewater typically involves physical treatment through primary clarification and biological treatment through activated sludge, oxidation ditches, and/or trickling filters followed by disinfection and disposal. Advanced secondary treatment includes secondary treatment of wastewater as well as conversion of nitrogen compounds through nitrification and/or denitrification, followed by disinfection. Tertiary treatment involves secondary treatment followed by media or membrane filtration and disinfection and may or may not include nitrification and denitrification. A breakdown of the level of treatment[1] provided by Central Valley POTWs with NPDES permits is presented in **Figure 1**. The majority of Central Valley POTWs with NPDES permits provide tertiary treatment. Several existing secondary or advanced secondary POTWs with NPDES permits are in the process of upgrading to provide tertiary treatment in the future.

[2] These categories generally follow the Central Valley POTW classifications from the CVCWA Methylmercury Control Study Progress Report (October 2015).

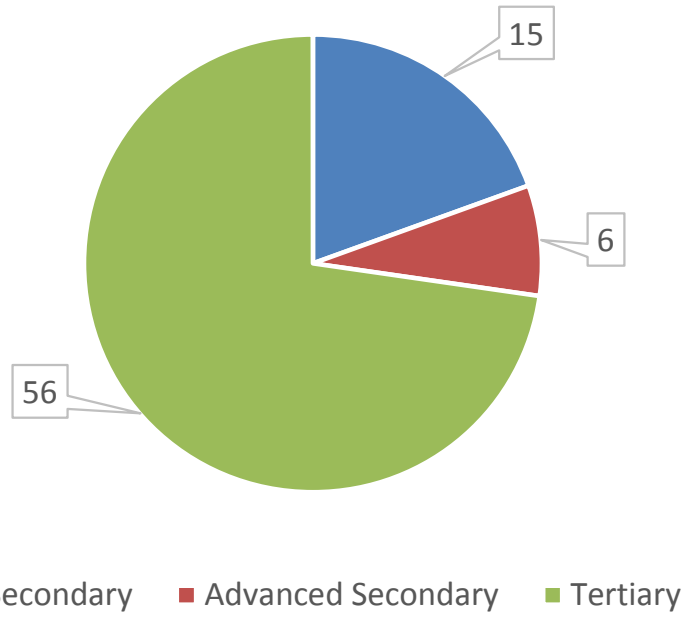

Figure 1. Number of Central Valley Publicly-Owned Treatment Works with NPDES Permits and Treatment Level.

2.2 WASTEWATER DISINFECTION METHODOLOGIES

Treated wastewater is disinfected prior to discharge to surface waters to provide human health protection to users of the water for recreation or drinking water supply. In the Central Valley, POTWs provide disinfection through either a chlorine-based or ultraviolet light system.

Chlorine-based disinfection systems require a dosage of chlorine-based chemicals that is added to the treated effluent prior to discharge to kill or render microorganisms inactive. Many POTWs that chlorinate treated effluent use sodium hypochlorite for chlorination followed by sodium bisulfite for dechlorination. Other POTWs use chlorine gas or chlorine dioxide for disinfection and sulfur dioxide gas for dechlorination. Advantages of chlorine-based disinfection include not needing wastewater treatment facilities beyond biological treatment (i.e., not needing media or membrane filtration) since effective chlorination is not dependent on near complete removal of particulates. Additionally, a residual of chlorine can help reduce the potential for bacterial regrowth or viral reactivation prior to discharge. Disadvantages of chlorine-based disinfection include human health and safety concerns associated with handling corrosive and toxic chemicals, production of disinfection byproducts (e.g., trihalomethanes), an increase in salinity of the effluent, and a potential risk to aquatic life if chlorine is not effectively removed from the final effluent through dechlorination prior to discharge.

For ultraviolet light disinfection, treated effluent, which has typically been filtered to reduce particulate concentrations, is passed through banks of high-intensity ultraviolet lights that kill or render microorganisms inactive. A reduced level of particulate matter in the final discharge is necessary for effective ultraviolet light disinfection as solids can shield microorganisms from exposure to ultraviolet light. Advantages of ultraviolet light disinfection include the absence of

need for corrosive and toxic chemicals in the disinfection process and avoidance of disinfection byproducts in the final effluent. Disadvantages of ultraviolet light disinfection are the capital and operational cost of media or membrane filtration and the higher energy requirement to operate ultraviolet lamps.

A breakdown of Central Valley POTWs utilizing chlorination- or ultraviolet light-based disinfection is presented in **Figure 2**. Several POTWs are currently in the process of changing their disinfection process from chlorination to ultraviolet light.

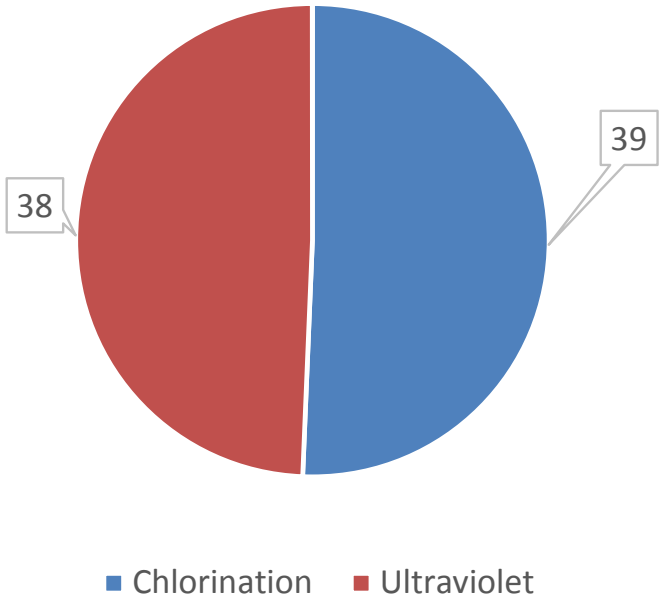

Figure 2. Number of Central Valley Publicly-Owned Treatment Works with NPDES Permits and Disinfection Methodology

This page is intentionally left blank.

Section 3. Chronic Toxicity Testing Requirements

This section provides general background on the purpose and requirements for chronic toxicity testing by Central Valley POTWs. These requirements include those currently contained in the *Water Quality Control Plan for the Sacramento River and San Joaquin River Basins, Fourth Edition* (Basin Plan) and NPDES permits, and those that are anticipated to be proposed in the State Water Board's Statewide Toxicity Policy.[3]

3.1 REGULATORY REQUIREMENTS

This section includes information on the federal and state regulatory requirements for chronic toxicity testing by POTWs.

3.1.1 Water Quality Control Plan for the Sacramento River and San Joaquin River Basins

The Clean Water Act prohibits the discharge of "pollutants in toxic amounts". Regulations to meet this narrative objective may be based on effluent limitations or water quality approaches that apply criteria for both chemical-specific parameters and whole effluent toxicity to protect water bodies and their beneficial uses.[4] USEPA's *Technical Support Document for Water Quality-based Toxics Control* (1996) states that biological assessments, as such techniques become available, should be integrated into water quality-based toxics control, thus creating a triad: whole effluent, chemical-specific, and biological assessments.

The Basin Plan, which was adopted by the Central Valley Water Board and has been effective since September 1998, includes a narrative toxicity objective that states the following:

> "*All waters shall be maintained free of toxic substances in concentrations that produce detrimental physiological responses in human, plant, animal, or aquatic life.*"

The Central Valley Water Board uses the *Policy for Implementation of Toxics Standards for Inland Surface Waters, Enclosed Bays, and Estuaries of California* (SIP, 2005) to implement the narrative toxicity objective from the Basin Plan through requirements for chronic toxicity testing in NPDES permits for Central Valley POTWs.

---

[3] California State Water Resources Control Board. *Summary of Proposed Toxicity Provisions* (April 2017). https://www.waterboards.ca.gov/water_issues/programs/state_implementation_policy/tx_ass_cntrl.shtml, Last accessed February 7, 2018.

[4] United States Environmental Protection Agency. *Toxicity Reduction Evaluation Guidance for Municipal Wastewater Treatment Plants* (EPA/833B-99/002). August 1999.

3.2 NPDES PERMIT REQUIREMENTS

As discussed above, the Central Valley Water Board requires chronic toxicity testing as one element of its program to meet the requirements of the Basin Plan (and Clean Water Act) in protecting aquatic life beneficial uses of surface waterbodies. These chronic toxicity testing requirements are generally found in Attachment E (Monitoring and Reporting Program) of Central Valley NPDES permits. The following sections provide general information on the requirements for chronic toxicity testing by Central Valley POTWs.

3.2.1 Routine Testing

In Central Valley NPDES permits, POTWs are typically required to conduct periodic chronic toxicity testing for three freshwater species: fathead minnow (*Pimephales promelas*), water flea (*Ceriodaphnia dubia*), and green alga (*Selenastrum capricornutum*, also known as *Raphidocelis subcapitata*). *P. promelas* is tested for 7 days for the growth and survival endpoints, *C. dubia* is tested for 6-8 days for survival and reproduction, and *S. capricornutum* is tested for 96 hours for growth. The testing frequency varies depending on the size of the POTW and ranges from once per NPDES permit term (e.g., once every five years) to monthly. Most Central Valley POTWs are required to conduct chronic toxicity testing either on an annual or quarterly basis during periods of discharge to surface receiving waters. The chronic toxicity testing frequency based on POTW size is presented in **Figure 3**. All sample collection and testing must adhere to USEPA's *Short-term Methods for Estimating Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms, Fourth Edition* (EPA/821-R-02-013, October 2002).

In NPDES permits adopted prior to the end of 2014, routine chronic toxicity testing of final effluent was required at different dilutions with receiving water and/or laboratory water (collectively, control water). Typical dilution series found in NPDES permits have been 100, 75, 50, 25, 12.5, and 0 percent (i.e., 100 percent control) or 100, 50, 25, 12.5, 6.25, and 0 percent of effluent mixed with control water. In NPDES permits adopted around the end of 2014 through 2016 for POTWs that had a chronic toxicity trigger of 1 TU_c, the routine chronic toxicity testing requirements typically did not included a dilution series (i.e., only testing of 100 percent effluent and 100 percent control water was required as part of routine monitoring). In these permits, dilution series testing is only required upon entering a TRE. NPDES permits adopted beginning in 2017 have reverted back to requiring a dilution series as part of the routine and accelerated chronic toxicity testing.

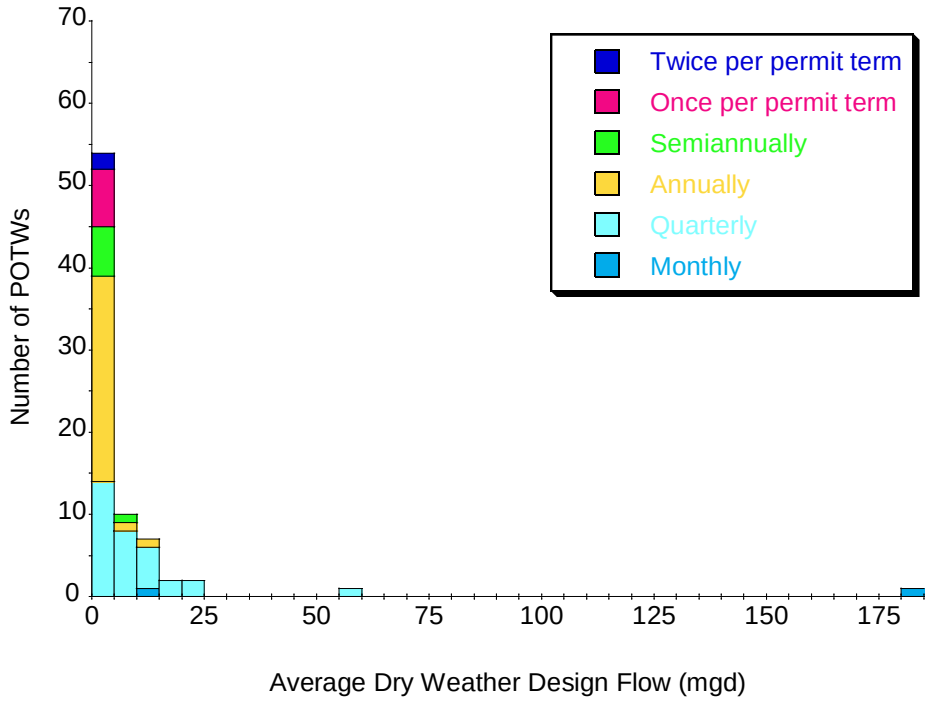

Figure 3. Required Chronic Toxicity Testing Frequency by Central Valley POTW Average Dry Weather Design Flow.

Because chronic toxicity testing utilizes live organisms, there is potential for chronic toxicity tests to be susceptible to variability and changes in the sensitivity of the test organisms and test failure. Variability in chronic toxicity testing is evaluated by testing laboratories by using percent minimum significant difference (PMSD) and in some cases (e.g., *S. capricornutum*) the coefficient of variation of the control treatment. NPDES permits require a concurrent reference toxicant test as a baseline test to document ongoing laboratory performance and to evaluate the sensitivity of the test organisms. Test failures are defined as either the effluent or reference toxicant test not meeting all test acceptability criteria. For instance, if an effluent is determined to be not toxic and the PMSD measured for the test exceeds the upper PMSD bound variability criteria, this renders the test to be too insensitive. In the event of a test failure, POTWs are required to re-sample and re-test.

Test results are currently required to be reported in terms of chronic toxicity units (TU_c). As of March 2017 (the end of the chronic toxicity data set for Central Valley POTWs evaluated for this study), Central Valley POTWs were required to calculate TU_c as 100 divided by the "No Observable Effect Concentration" (NOEC). The NOEC is statistically determined as the lowest effluent test concentration in which there was no observed significant effect.

Most Central Valley POTWs chronic toxicity triggers do not consider dilution and therefore have a numeric toxicity monitoring trigger of 1 TU_c on the basis of the NOEC as determined by

hypothesis testing. USEPA requirements for NPDES permits may also require that chronic toxicity test reports indicate the concentration at which 50 percent of the test organisms exhibit mortality (LC_{50}); concentration at which 25 percent of the test organisms exhibit an effect or inhibition (EC_{25} or IC_{25}); or concentration at which 50 percent of the test organisms exhibit inhibition for the test endpoint (IC_{50}). Exceedance of a 1 TU_c trigger requires follow-up accelerated testing, which is discussed in the following section.

The Central Valley Water Board has the discretion to provide a higher numeric toxicity monitoring trigger if receiving water dilution is considered. As a result, several Central Valley POTWs have higher numeric toxicity monitoring triggers (e.g., 4 TU_c, 8 TU_c).

3.2.2 Accelerated Testing

Accelerated testing is required if a toxicity test indicates an exceedance of the chronic toxicity trigger, and typically consists of four chronic toxicity tests conducted once every two weeks. If none of those four accelerated chronic toxicity tests exceed the chronic toxicity trigger, the POTW can cease accelerated testing and resume routing testing in accordance with the testing frequency required in its NPDES permit. However, if any of the four accelerated tests indicates an exceedance of the chronic toxicity trigger, the POTW is required to initiate a TRE, which a process intended to identify the cause(s) and source(s) of the toxicity and is discussed in greater detail in the following section. NPDES permits allow dischargers who readily identify the cause of toxicity during accelerated testing to take corrective actions and reinitiate accelerated testing to verify that there is no longer an indication of toxicity in the effluent.

Beginning with NPDES permits adopted in December 2017, the Central Valley Water Board made modifications to its approach for requiring accelerated testing. POTWs were required to conduct accelerated testing only if both the chronic toxicity trigger was exceeded and the results in the 100 percent effluent were at least 25 percent less than the control. This approach is expected to help resolve some of the resource-intensive follow-up issues with potential low-level chronic toxicity that are discussed in this report.

3.2.3 Toxicity Reduction Evaluations/Toxicity Identification Evaluations

As noted above, a TRE is typically initiated when effluent toxicity is observed to be persistent, which is most often demonstrated when repeated chronic toxicity tests indicate the presence of toxicity during accelerated testing. The goal of a TRE is to reduce or eliminate effluent toxicity. Specifically, a TRE is:

> *"A site-specific study conducted in a step-wise process designed to identify the causative agent(s) of effluent toxicity, isolate the sources of toxicity, evaluate the effectiveness of toxicity control options, and then confirm the reduction in effluent toxicity."* (USEPA 1991)

In *Toxicity Reduction Evaluation Guidance for Municipal Wastewater Treatment Plants* (USEPA 1999), USEPA presents an acceptable protocol for implementing a TRE. Typically, this is the protocol that is most often followed or referenced by NPDES permittees and permit writers when a TRE is initiated. The TRE protocol is generally divided into six component parts, or tiers, including:

1. Information and Data Acquisition (Tier 1)
2. Facility Performance Evaluation (Tier 2)
3. Toxicity Identification Evaluation (Tier 3)
4. Toxicity Source Evaluation (Tier 4)
5. Toxicity Control Evaluation (Tier 5)
6. Toxicity Control Implementation (Tier 6)

This TRE protocol is schematically presented in **Figure 4**, and each component tier of the TRE protocol is described briefly below. It should be noted that while the protocol is presented in a linear fashion (i.e., sequentially numbered tiers and a linear flow chart), a TRE can be successfully concluded at any stage of the process, and some tiers can be omitted, conducted concurrently, or implemented out of sequence. Moreover, the TRE protocol is not a mandate, and activities other than those described in the protocol that lead to effluent toxicity resolution may also be considered. As clarified by USEPA in *Clarifications Regarding Toxicity Reduction and Identification Evaluations in the National Pollutant Discharge Elimination System Program* (USEPA 2001):

> "*Any activities that result in consistently reducing toxicity to an acceptable level may be considered TRE activities.*"

Information and Data Acquisition (Tier 1): Under Tier 1, all relevant data pertaining to the indication of toxicity are compiled. These data may include, but are not limited to a) effluent toxicity data, b) influent and effluent monitoring data, c) process monitoring data, d) pretreatment monitoring data and industrial waste surveys, and e) facility and collection system operation and maintenance logs.

Facility Performance Evaluation (Tier 2): Under Tier 2, compiled information and data are reviewed and evaluated with an emphasis on determining whether activities at the treatment facility or within the collection system (including actions by significant industrial users) may have caused or contributed to the indication of toxicity. Information gained in this tier may result in identification and eventual elimination of confirmed effluent toxicity, or may provide essential information to focus or target subsequent tiers, especially TIE and Toxicity Source Evaluation tiers.

Toxicity Identification Evaluation (Tier 3): Under Tier 3, effluent is subject to a forensic evaluation in an effort to identify the contaminant(s) responsible for the indicated toxicity. Specifically, a TIE is a laboratory-based study where effluent can be chemically and/or physically manipulated with the goal of the manipulation being the removal, recovery, and/or amplification of toxicity in a sample. Patterns of toxicity removal, recovery, and/or amplification provide clues as to the contaminant(s) responsible for toxicity. A TIE may be performed in phases (i.e., Phase I, II, and III TIEs) where each subsequent phase of the TIE process is intended to provide further resolution and confirmation of the specific contaminant(s) responsible for toxicity. The specific degree of contaminant resolution and/or confirmation required in a TIE is often dictated by ease of mitigating effluent toxicity. For example, in cases where methods of toxicity control include large capital improvements (e.g., new or upgraded treatment/pretreatment facilities), definitive confirmation of the identified toxicant would be desired, thus the TIE methodologies employed may reach Phase III. In some cases, sufficient

information is obtained by an initial Phase I TIE in order for the source of toxicity to be identified and/or controlled.

Toxicity Source Evaluation (Tier 4): Under Tier 4, focus is given to identify "upstream" sources of toxicity, possibly with no particular emphasis given to the specific toxicant(s) involved, but ultimately with the goal of identifying and controlling toxicity at its source. Activities associated with source evaluation may include pre- and post-treatment process chronic toxicity testing or process side-stream testing, or may include collection system or significant industrial user discharge chronic toxicity and/or analytical chemistry testing. Information gathered during the Facility Performance Evaluation (Tier 2) may lead directly to Toxicity Source Evaluation with no need for performing a TIE (Tier 3).

Toxicity Control Evaluation (Tier 5): Under Tier 5, toxicity control options are evaluated and selected. This evaluation may consider new or modified industrial pretreatment requirements or treatment facility modifications. Moreover, this evaluation may include treatability studies aimed at developing methods for optimizing existing treatment processes or identifying new treatment facilities or processes, including pilot testing of new or proprietary treatment systems or chemical additives, with chronic toxicity testing as confirmation. Overall, the objective of this step is to select the most feasible, cost-effective option for effluent toxicity reduction considering all technical and cost factors.

Toxicity Control Implementation (Tier 6): Under Tier 6, the selected toxicity control solution, if identified, is implemented. Depending on the findings of the TRE, implementation may involve relatively minor changes such as modifying treatment facility operating procedures or more complex modifications such as expanding the source control or pretreatment program or designing and constructing new treatment facilities or processes.

Conclusion of a TRE in the Absence of Identifying a Cause for Toxicity: In practice, it is not uncommon for a TRE to be concluded without definitively identifying the cause or reason for observed toxicity. This information is further developed in **Section 4** of this report. Particularly in the case of a single unrepeated incident, the indication of effluent toxicity may cease before a cause can be identified and/or before a control solution can be implemented. In these cases, it is important to reflect upon the principal objective of the TRE, that being the reduction or elimination of effluent toxicity. The consistent absence of toxicity, in light of an aggressive pursuit of the cause, but where the cause remains unknown, is still a sufficient basis to conclude a TRE. As clarified by USEPA (2001):

> *"If the toxic event is an isolated incident, demonstrated by lack of toxicity in retesting over an extended period (e.g., during the next three months of accelerated testing), the TRE may be discontinued if the regulatory authority is satisfied that an additional occurrence is unlikely."*

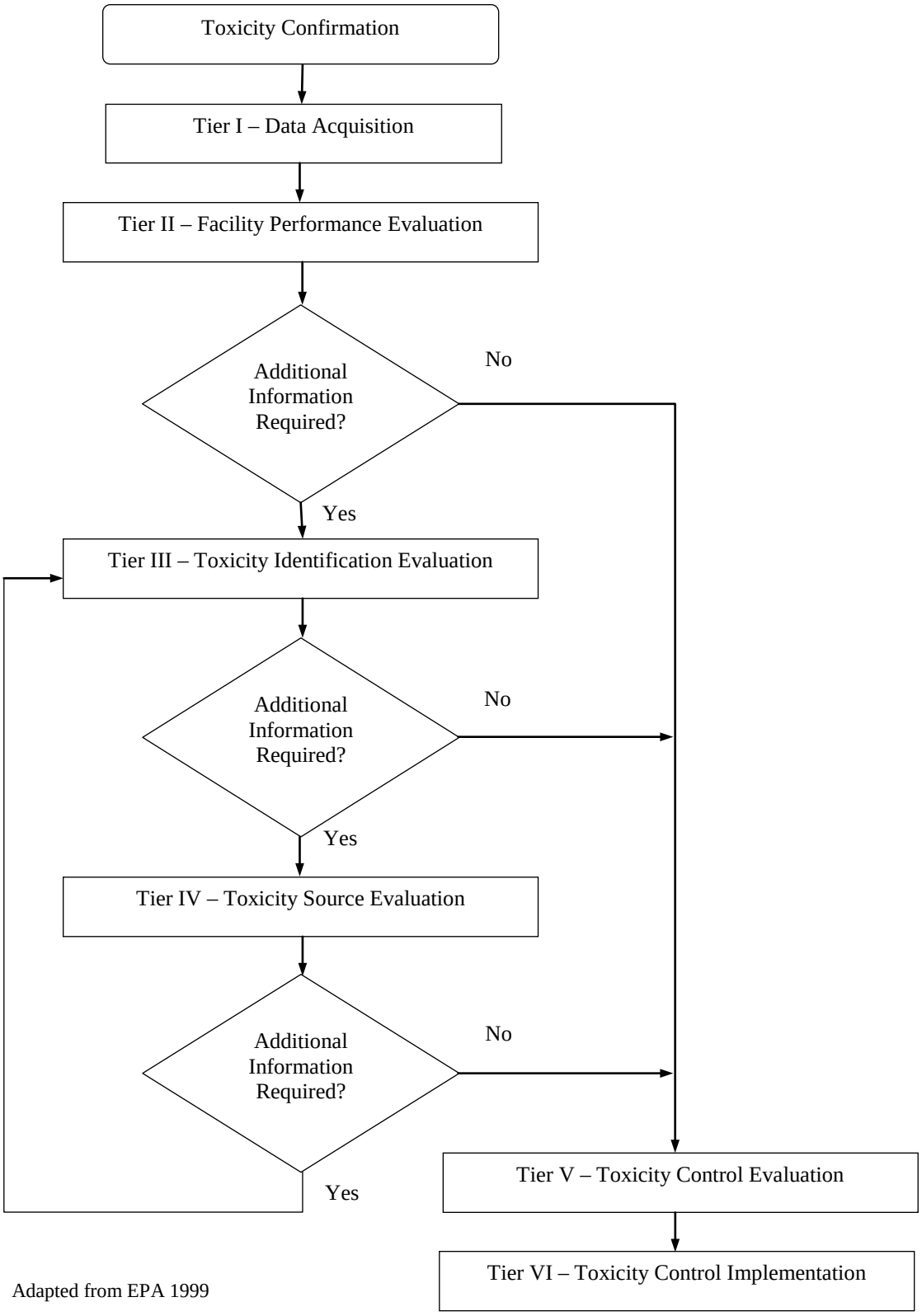

Figure 4. Toxicity Reduction Evaluation Protocol.

3.3 PROPOSED STATEWIDE TOXICITY POLICY

The State Water Board is currently in the process of developing a Statewide Toxicity Policy to establish water quality objectives for aquatic toxicity and a statistical approach for assessing toxicity in POTW and other effluents and receiving waters. The proposed policy would be included in the statewide *Water Quality Control Plan for Inland Surface Waters, Enclosed Bays, and Estuaries* and would supersede current requirements in Basin Plans. The proposed policy would establish numeric water quality objectives for both acute and chronic toxicity and establish a program of implementation to control toxicity. Under the most recent publicly-available proposal[5], attainment of the water quality objective would be demonstrated by rejecting the null hypothesis, which states that "the ambient receiving water is toxic because the test organism adverse response in ambient receiving water sample is significantly different from the test organism response in the control water sample." The statistical analysis will be conducted using the Test of Significant Toxicity (TST) (USEPA 2010).

Some key proposed provisions contained in the 2012 public draft of the Statewide Toxicity Policy included provisions that all POTWs over 5 mgd would automatically be assumed to have reasonable potential where those less than 5 mgd a reasonable potential analysis would be conducted to determine if effluent limitations were required. For those with reasonable potential, numeric maximum daily effluent limitations (MDELs) and median monthly effluent limitations (MMELs) would be included in NPDES permits. Additionally, in many cases, the frequency of routine monitoring would be increased. The 2012 proposed policy would require TREs to be initiated if there are two or more exceedances of a chronic toxicity effluent limitation in a two-month period.

The fact that, under the proposed Statewide Toxicity Policy, MDELs and MMELs for chronic toxicity are being proposed with thresholds applicable to the difference between IWC and control toxicity tests with sub-lethal endpoints (e.g., 50 percent effect; no statistically significant effect) warrants an assessment of these metrics for chronic toxicity test results among Central Valley POTWs. This represents a shift from the historic focus on TU_c metrics in chronic toxicity test results, and warrants an assessment of the "effect" metric (percent difference from control) in evaluating the values at which multiple laboratories agree on toxicity and factors related to laboratory variability, and the relationship between these effect metrics and measurable effects in the receiving water are expected to occur.

As of September 2018, the State Water Board has not yet re-released a public review draft of the Statewide Toxicity Policy. The State Water Board updates its website with an anticipated schedule for release, public comment, and consideration. However, these dates are tentative and subject to change.

[5] California State Water Resources Control Board. *Summary of Proposed Toxicity Provisions* (April 2017). https://www.waterboards.ca.gov/water_issues/programs/state_implementation_policy/tx_ass_cntrl.shtml, Last accessed February 7, 2018.

Section 4. Characterization of Chronic Toxicity Test Results in Central Valley POTWs

As discussed in **Section 1**, one of the goals of this study is to understand and characterize the level at which Central Valley POTWs observed exceedances of the chronic toxicity trigger and low-level chronic toxicity and the methods and steps that are taken by POTWs to identify and resolve these issues. This section provides a review of chronic toxicity test results from Central Valley POTWs, the analyses of those test results to understand the nature of chronic toxicity trigger exceedances, and the approaches taken by POTWs to resolve chronic toxicity issues.

4.1 DATA SOURCES

The State Water Board maintains the California Integrated Water Quality System (CIWQS), which is a central data and reporting repository for POTWs throughout California. POTWs operating under NPDES permits submit periodic reports to CIWQS that primarily consist of water quality data and test results for effluent discharges, but may also contain other information or data (e.g., annual operations reports, periodic pretreatment reports, toxicity reports) required by individual NPDES permits.

For this analysis, chronic toxicity test reports and data were collected for the period of January 2011 to March 2017 by accessing the CIWQS database. Most Central Valley POTWs with NPDES permits upload chronic toxicity test reports, which are prepared by toxicity testing laboratories. Most reports are uploaded to CIWQS as attachments, with test results summarized through data entry and maintained in a report table on the CIWQS website. These chronic toxicity test reports are downloadable by the public. Upon reviewing the readily available chronic toxicity information and data from CIWQS, additional targeted data requests were made to Central Valley Water Board staff and to some individual POTWs to obtain additional chronic toxicity reports to fill in data gaps.

The following were key observations related to gathering chronic toxicity data and test reports used for this study:

- The purpose of obtaining full chronic toxicity test reports instead of only compiling TU_c results, which are readily available in CIWQS, was to document the level of "effect" observed in the effluent relative to the control (e.g., *C. dubia* reproductive inhibition, *S. capricornutum* cell growth inhibition). Ideally, this will provide information to allow future evaluation of effects levels at which follow-up TRE efforts were able to successfully identify the cause(s) of indicated toxicity.
- Chronic toxicity test reports were not consistently uploaded to CIWQS prior to about 2011 or 2012. Subsequent to this period, until March 2017 (the end date for the data set used in this analysis), most routine chronic toxicity test reports conducted by larger POTWs were uploaded to CIWQS and were available for this analysis.
- Smaller POTWs were not required or were on a different time schedule to submit electronic data to CIWQS. As such, chronic toxicity test reports and data for smaller Central Valley POTWs were not readily available.

- Chronic toxicity test reports associated with accelerated and TRE/TIE testing were not always available for download from CIWQS. However, those chronic toxicity test reports made available through CIWQS, by the Central Valley Water Board, or by individual POTWs were included in this analysis.

Of the 77 Central Valley POTWs identified in **Figures 1** and **2**, chronic toxicity data from 66 Central Valley POTWs were included in this analysis. While the special study consultant team made a concerted effort in collaboration with CVWCA and the Central Valley Water Board to obtain as many of the chronic toxicity test reports (i.e., routine chronic toxicity testing, accelerated testing, TRE/TIE testing) as possible, it is acknowledged that it was not possible to obtain all chronic toxicity test reports from all Central Valley POTWs during the data period evaluated. A comparison of total Central Valley POTWs to POTWs that were included in this study is presented in **Table 1**. The POTWs for which chronic toxicity test data were obtained for this study are believed to be representative of all 77 Central Valley POTWs.

Table 1. Comparison of POTWs in Toxicity Special Study to All Central Valley POTWs.

	Total Central Valley POTWs	Number of POTWs in Study	Total Number of Chronic Toxicity Test Reports (1)
Treatment Type			
Secondary	15 (19%)	14 (21%)	620 (21%)
Advanced Secondary	6 (8%)	5 (8%)	145 (5%)
Tertiary	56 (73%)	47 (71%)	2,187 (74%)
Disinfection Process			
Chlorination (2)	39 (51%)	31 (47%)	1,144 (39%)
Ultraviolet (2)	38 (49%)	35 (53%)	1,808 (61%)

(1) Chronic toxicity test reports include routine chronic toxicity testing, accelerated testing, and TRE/TIE-related testing.
(2) Nevada County Sanitation District No. 1 Lake Wildwood Wastewater Treatment Plant converted from chlorination to ultraviolet light disinfection in 2013. The data analyses considers the disinfection method utilized at the time of the chronic toxicity test and groups information appropriately. For this table, this facility is considered an ultraviolet light disinfection facility.

4.2 ROUTINE CHRONIC TOXICITY TEST RESULTS ANALYSES

Analyses were conducted on routine chronic toxicity test results to characterize chronic toxicity test results from Central Valley POTWs. The purpose of these analyses was to identify trends and other factors that may potentially influence chronic toxicity testing and provide guidance for developing approaches to address observed chronic toxicity in effluent discharges from Central Valley POTWs. The subsequent sections summarize the analyses conducted for the following factors:

- Baseline characterization;
- POTW treatment level;
- Nitrogen treatment; and
- Disinfection methodology.

4.2.1 Baseline Characterization

As discussed in the previous section, chronic toxicity test reports were obtained for 66 Central Valley POTWs. Nearly 1,000 routine chronic toxicity tests (i.e., testing conducted according to NPDES permit monitoring frequencies) were obtained for each of the three test species (e.g., *P. promelas*, *C. dubia*, and *S. capricornutum*) and evaluated in this analysis. In distinguishing routine chronic toxicity test data and other chronic toxicity test data (e.g., accelerated testing, TRE testing), routine chronic toxicity tests were generally identified by when the sampling and toxicity test occurred in relation to the required NPDES permit monitoring frequency for each POTW. There were instances where routine chronic toxicity testing overlapped with accelerated testing and/or TRE/TIE testing. For these situations, the tests (including baseline tests for TRE/TIE testing) were classified as routine chronic toxicity tests for the purpose of this analysis. (Chronic toxicity tests identified as accelerated testing, TRE/TIE testing, or ones conducted at a frequency greater than the required NPDES permit monitoring frequency were excluded from the baseline characterization.) In reviewing the routine chronic toxicity testing frequency requirements, it is estimated that the data set utilized in this analysis is missing approximately 200 routine chronic toxicity tests based on the expected NPDES permit testing frequency for the POTWs included in this study. This means that approximately 80 to 85 percent of the expected total routine chronic toxicity tests between January 2011 and March 2017 are included in the data set for this analysis. These test results were used to evaluate various factors such as treatment type, disinfection methodology, and seasonality.

It was critical to note that some Central Valley POTWs have chronic numeric toxicity triggers that considered receiving water dilution, which resulted in increasing triggers from 1 TU_c up to a maximum of 16 TU_c. Of the 66 Central Valley POTWs included in this study, 7 POTWs have a chronic toxicity trigger higher than 1 TU_c. With the exception of the City of Woodland, which receives a chronic toxicity trigger of 2 TU_c for *S. capricornutum* only, the remaining 58 Central Valley POTWs (59 Central Valley POTWs including the City of Woodland for *P. promelas* and *C. dubia*) evaluated in this study have a chronic toxicity trigger of 1 TU_c.

An overall summary of chronic toxicity trigger exceedances for Central Valley POTWs during routine chronic toxicity testing for each test endpoint is presented in **Table 2** along with the number of POTWs impacted by chronic toxicity trigger exceedances.

Table 2. Central Valley POTWs Chronic Toxicity Trigger Exceedances, January 2011 to March 2017.

Test Organism/Endpoint	Total Number of Chronic Toxicity Tests	Number of Chronic Toxicity Trigger Exceedances (%)	Number and Percent of Central Valley POTWs Impacted
Central Valley POTWs with Chronic Toxicity Trigger = 1 TU_c (1)			
Pimephales promelas (survival)	832	4 (0.5%)	2 (3.4%)
Pimephales promelas (growth)	834	20 (2.4%)	15 (25.4%)
Ceriodaphnia dubia (survival)	818	9 (1.1%)	7 (11.8%)
Ceriodaphnia dubia (reproduction)	820	131 (16.0%)	38 (64.4%)
Selenastrum capricornutum (growth)	835	76 (9.1%)	29 (49.2%)
Central Valley POTWs with Chronic Toxicity Trigger > 1 TU_c (2)			
Pimephales promelas (survival)	128	2 (1.6%)	1 (14.3%)
Pimephales promelas (growth)	128	2 (1.6%)	1 (14.3%)
Ceriodaphnia dubia (survival)	137	0 (0.0%)	0 (0.0%)
Ceriodaphnia dubia (reproduction)	138	17 (12.3%)	3 (42.9%)
Selenastrum capricornutum (growth)	152	2 (1.3%)	1 (12.5%)

(1) There are 59 Central Valley POTWs that have a chronic toxicity trigger of 1 TU_c for all test organisms. Prior to December 1, 2014, the City of Woodland had a chronic toxicity trigger of 1 TU_c for *S. capricornutum.* Chronic toxicity data collected for the City of Woodland prior to this date were included in this subset of data.

(2) There are 7 Central Valley POTWs that have a chronic toxicity trigger of greater than 1 TU_c for all test organisms. After December 1, 2014, the City of Woodland had a chronic toxicity trigger of 2 TU_c for *S. capricornutum.* Chronic toxicity data collected for the City of Woodland after this date were included in this subset of data.

Most chronic toxicity trigger exceedances for Central Valley POTWs were observed for *C. dubia* reproduction and *S. capricornutum* growth for POTWs that have a chronic toxicity trigger of 1 TU_c. Approximately two-thirds of the POTWs experienced at least one chronic toxicity trigger exceedance for *C. dubia* reproduction and half of the POTWs observed at least one exceedance for *S. capricornutum* growth during the time period assessed for this study. For Central Valley POTWs that have a chronic toxicity trigger of greater than 1 TU_c, the *C. dubia* reproduction test endpoint was exceeded more frequently than the other chronic toxicity endpoints, and affected approximately half of the POTWs in this category. Central Valley POTWs that have dilution (and higher chronic toxicity triggers) had significantly fewer exceedances compared to POTWs that have a chronic toxicity trigger of 1 TU_c. Exceedances of the chronic toxicity trigger for *P. promelas* survival and growth and *C. dubia* survival were infrequently observed by Central Valley POTWs. It should be noted that the *P. promelas* growth and *C. dubia* reproduction endpoints are dependent on the survival endpoints for those species such that mortality of the test organisms affects the outcome of the other endpoints. Because the survival endpoint in chronic toxicity tests does not appear to be a significant issue for *P. promelas* or *C. dubia* in the Central Valley, these test endpoints were precluded from further analysis in this study.

Through the time period of the data set, permitting approaches to chronic toxicity evaluations were changed to address information learned from chronic toxicity testing findings and effectiveness and impacts of the testing and TREs, including costs associated with conducting follow-up investigation of toxicity trigger exceedances. This made it difficult to evaluate certain

trends associated with chronic toxicity trigger exceedances using TU_c. Examples of the changes in permitting approaches include the following:

- The control water requirements for chronic toxicity testing differed among POTW NPDES permits. Older NPDES permits typically specified the control water to which effluent results were compared. More recent NPDES permits allow the POTW to use either laboratory or receiving water for the control water.
- The requirement for a dilution series for chronic toxicity testing changed during the data period evaluated.
    - In NPDES permits adopted prior to the end of 2014, routine chronic toxicity tests required a dilution series with either laboratory or receiving water.
    - In NPDES permits adopted from the end of 2014 through 2016, POTWs with a numeric toxicity monitoring trigger of 1 TU_c were not required to conduct chronic toxicity testing with a dilution series unless it was undergoing a TRE.
    - In NPDES permits adopted since 2017, routine chronic toxicity tests required a dilution series with either laboratory or receiving water.

The following sections provide general summaries of chronic toxicity test results for Central Valley POTWs between January 2011 and March 2017.

4.2.1.1 Chronic Toxicity Trigger Exceedances by Control Water

Generally in this analysis of all data, it was not possible to consistently and definitively determine the control water (e.g., receiving or laboratory water) used for chronic toxicity testing. This limited the ability to determine whether the control water affected the chronic toxicity test results. This information was inconsistently available in chronic toxicity test reports because multiple testing laboratories were used and reporting procedures between laboratories varied.

However, in reviewing the subset of chronic toxicity data in which the toxicity trigger was exceeded, it was possible to identify the control water that was used for the comparison of effluent that resulted in the determination that the chronic toxicity trigger was exceeded. The control water utilized to evaluate POTW effluent for *C. dubia* reproduction that resulted in an exceedance of the chronic toxicity trigger is presented in **Figure 5**. Similarly, the control utilized to evaluate POTW effluent for *S. capricornutum* growth that resulted in an exceedance of the chronic toxicity trigger is presented in **Figure 6**.

For Central Valley POTWs with a 1 TU_c chronic toxicity trigger, most exceedances of the chronic toxicity trigger for *C. dubia* reproduction occurred when laboratory water was used as the control. However, exceedances of the chronic toxicity trigger for *S. capricornutum* growth for these POTWs occurred equally between receiving water and laboratory water controls. One concern that has been raised, and is discussed further in **Section 5**, is that the use of receiving water for *S. capricornutum* growth may have biostimulatory effects (e.g., higher concentrations of nutrients or other characteristics that may promote test organism health in this control water). This test variability may result in observed toxicity in the effluent when compared to the biostimulatory control.

Most Central Valley POTWs that have a chronic toxicity trigger of greater than 1 TU_c use the receiving water as the control.

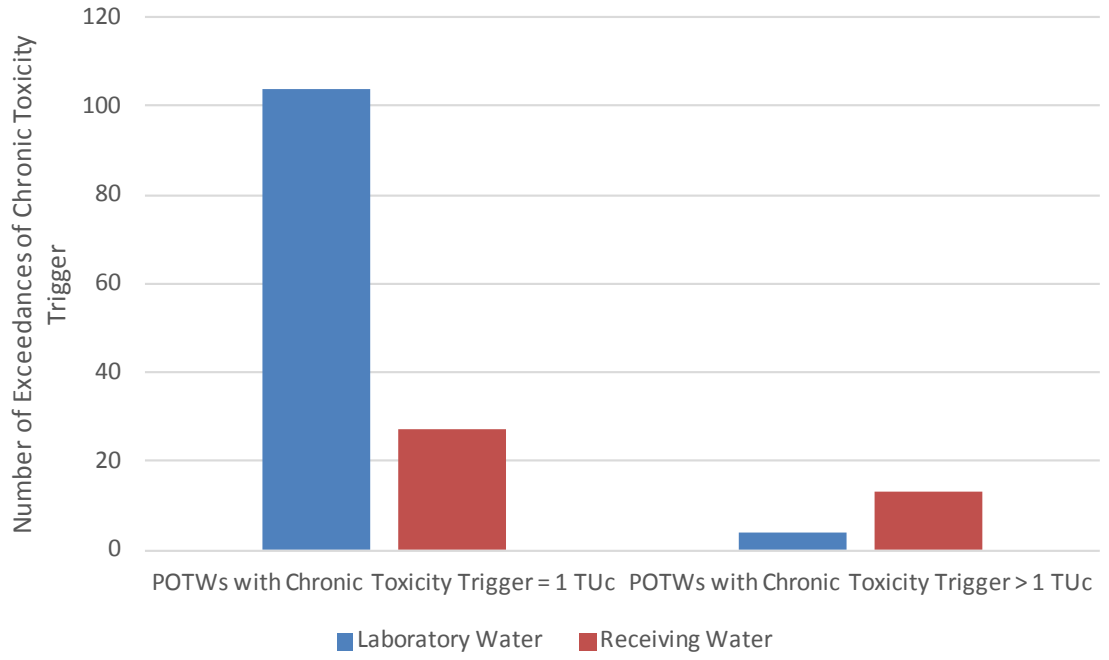

Figure 5. Control Water Used for *Ceriodaphnia dubia* Reproduction in Chronic Toxicity Trigger Exceedances.

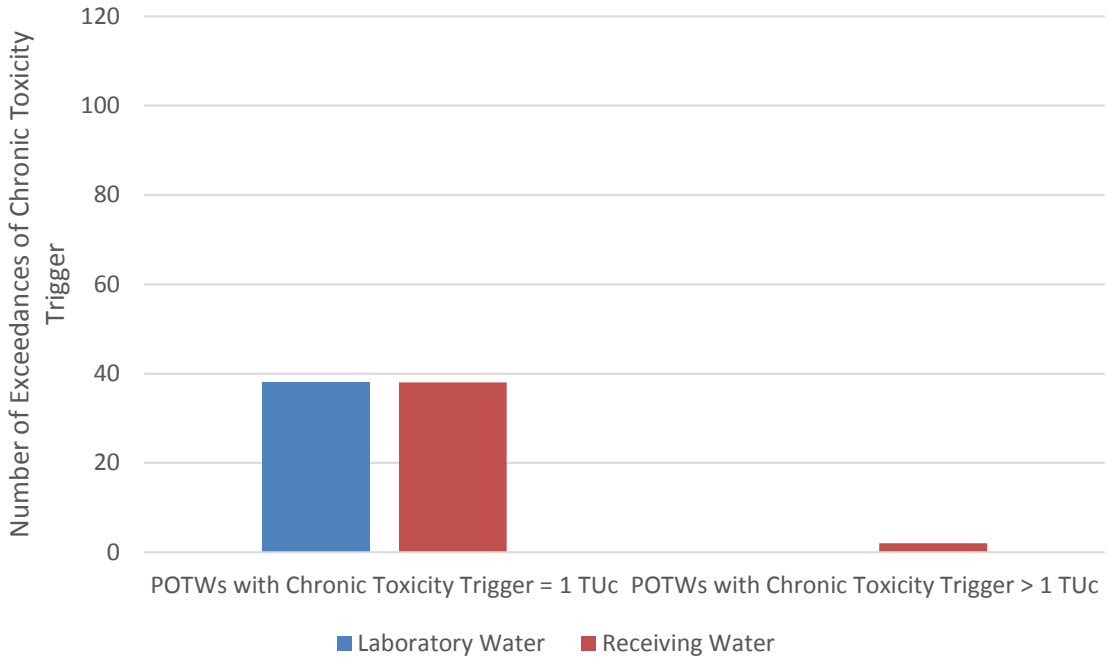

Figure 6. Control Water Used for *Selenastrum capricornutum* Growth in Chronic Toxicity Trigger Exceedances.

4.2.1.2 *Magnitude of Chronic Toxicity Trigger Exceedances using TU_c*

As stated previously, NPDES permits adopted by the Central Valley Water Board from the end of 2014 through 2016 only required chronic toxicity testing of 100 percent effluent and the control water(s). As a result, chronic toxicity test results were often reported as greater than 1 TU_c when chronic toxicity was observed in the effluent sample. This revised approach for testing resulted in data depicted as greater than 1 TU_c, which confounds attempts to define or determine whether the results were "low-level" or not. The levels of reported chronic toxicity trigger exceedance for *C. dubia* reproduction and *S. capricornutum* growth and exceedance frequencies are presented in **Figure 7** and **Figure 8**, respectively, for the available data for the entire study period.

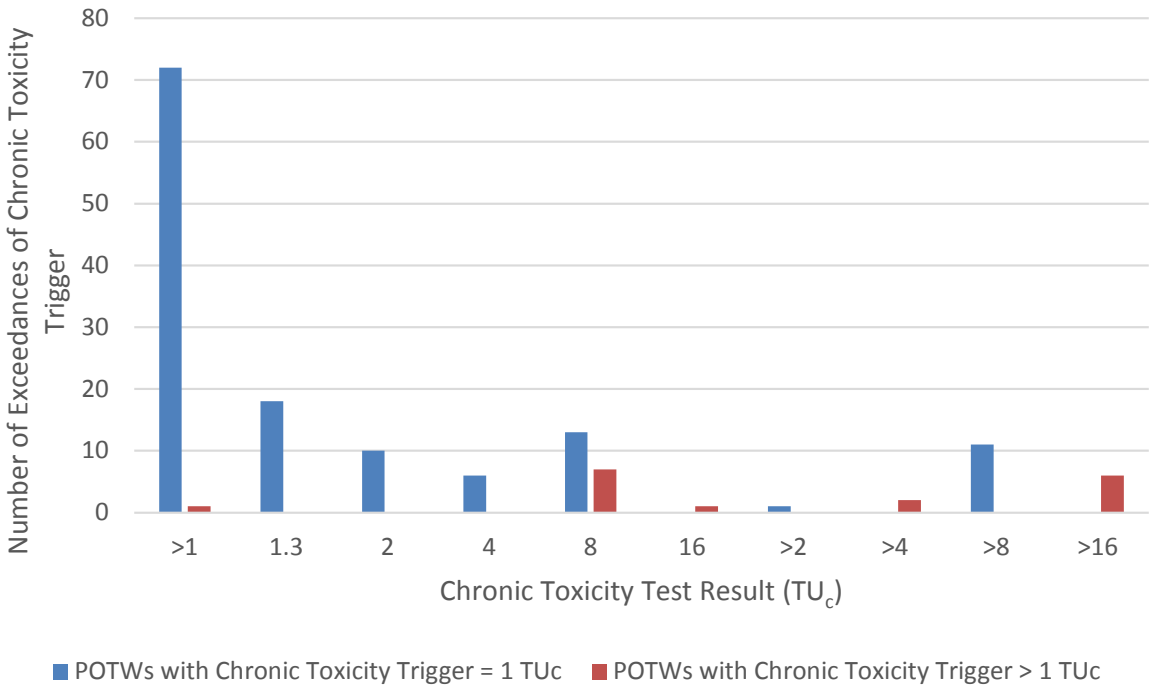

Figure 7. Chronic Toxicity Trigger Exceedances for *Ceriodaphnia dubia* Reproduction for Central Valley POTWs, January 2011-March 2017.

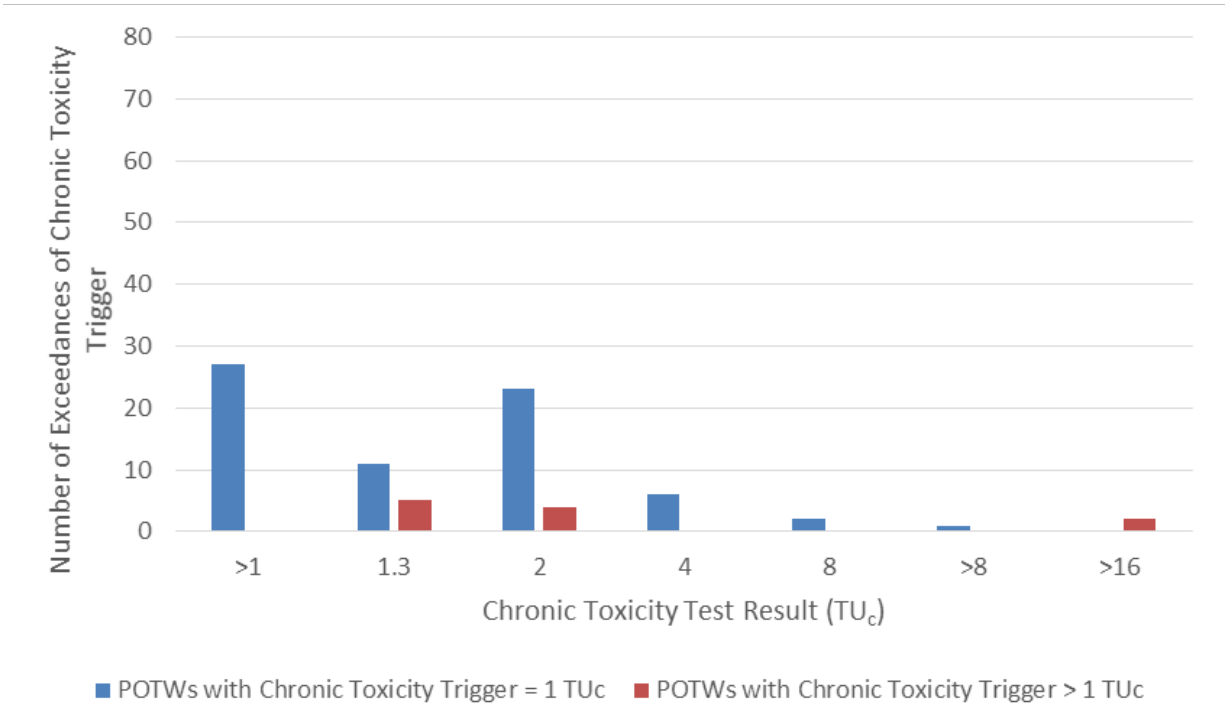

Figure 8. Chronic Toxicity Trigger Exceedances for *Selenastrum capricornutum* Growth for Central Valley POTWs, January 2011-March 2017.

Setting aside the reported chronic toxicity trigger results of greater than 1 TU_c, most of the remaining chronic toxicity trigger exceedances for *C. dubia* reproduction and *S. capricornutum* growth were 1.3 or 2 TU_c, which means that, in those cases, toxicity was generally only observed in the 100 percent effluent sample, and was not present when the sample was diluted by 25 or 50 percent, respectively.

4.2.1.3 Temporal Assessment

Potential temporal impacts were evaluated to determine if there was any indication that this could be a factor that can affect chronic toxicity tests. Breakdowns of chronic toxicity tests taken each year and the number of toxicity trigger exceedances for *C. dubia* and *S. capricornutum* are presented in **Table 3** and **Table 4**, respectively.

Table 3. Chronic Toxicity Trigger Exceedances for *Ceriodaphnia dubia* Reproduction for Central Valley POTWs by Year.

Year	Total Number of Chronic Toxicity Tests	Number of Chronic Toxicity Trigger Exceedances (%)	Number of Central Valley POTWs Impacted (%)
Central Valley POTWs with Chronic Toxicity Trigger = 1 TU_c (1)			
2011	118	9 (7.6%)	7 (11.9%)
2012	142	16 (11.3%)	11 (18.6%)
2013	136	20 (14.7%)	16 (27.1%)
2014	137	20 (14.6%)	15 (25.4%)
2015	132	27 (20.5%)	19 (32.2%)
2016	118	29 (24.6%)	21 (35.6%)
Jan-Mar 2017	37	10 (27.0%)	9 (15.3%)
Central Valley POTWs with Chronic Toxicity Trigger > 1 TU_c (2)			
2011	18	0 (0.0%)	0 (0.0%)
2012	30	0 (0.0%)	0 (0.0%)
2013	26	3 (11.5%)	1 (14.3%)
2014	21	3 (14.3%)	2 (28.6%)
2015	19	6 (31.6%)	2 (28.6%)
2016	18	3 (16.7%)	3 (42.9%)
Jan-Mar 2017	6	2 (33.3%)	2 (28.6%)

(1) There are 59 Central Valley POTWs that have a chronic toxicity trigger of 1 TU_c.
(2) There are 7 Central Valley POTWs that have a chronic toxicity trigger of greater than 1 TU_c.

Table 4. Chronic Toxicity Trigger Exceedances for *Selenastrum capricornutum* Growth for Central Valley POTWs by Year.

Year	Total Number of Chronic Toxicity Tests	Number of Chronic Toxicity Trigger Exceedances (%)	Number of Central Valley POTWs Impacted (%)
Central Valley POTWs with Chronic Toxicity Trigger = 1 TU_c (1)			
2011	121	8 (6.6%)	5 (8.5%)
2012	139	12 (8.6%)	9 (15.3%)
2013	144	10 (6.9%)	7 (11.9%)
2014	140	11 (7.9%)	8 (13.6%)
2015	133	18 (13.5%)	12 (20.7%)
2016	123	14 (11.4%)	11 (19.0%)
Jan-Mar 2017	35	3 (8.6%)	3 (5.2%)
Central Valley POTWs with Chronic Toxicity Trigger > 1 TU_c (2)			
2011	18	0 (0.0%)	0 (0.0%)
2012	30	2 (6.7%)	1 (14.2%)
2013	25	0 (0.0%)	0 (0.0%)
2014	21	0 (0.0%)	0 (0.0%)
2015	25	0 (0.0%)	0 (0.0%)
2016	26	0 (0.0%)	0 (0.0%)
Jan-Mar 2017	7	0 (0.0%)	0 (0.0%)

(1) There are 58 Central Valley POTWs that have a chronic toxicity trigger of 1 TU_c for *S. capricornutum.* Before December 1, 2014, the City of Woodland had a chronic toxicity trigger of 1 TU_c for *S. capricornutum.* Chronic toxicity data collected for the City of Woodland prior to this date were included in this subset of data.

(2) There are 7 Central Valley POTWs that have a chronic toxicity trigger of greater than 1 TU_c for *S. capricornutum*. After December 1, 2014, the City of Woodland has a chronic toxicity trigger of 2 TU_c for *S. capricornutum*. Chronic toxicity data collected for the City of Woodland after this date were included in this subset of data.

In a year to year analysis of chronic toxicity trigger exceedances for *C. dubia* reproduction, there appeared to be an increasing number of exceedances as a percentage of the number of chronic toxicity test performed as well as in the number of POTWs impacted for Central Valley POTWs that have a chronic toxicity trigger of 1 TU_c. For *S. capricornutum* growth, the number of chronic toxicity trigger exceedances for Central Valley POTWs with a chronic toxicity trigger of 1 TU_c is relatively constant with the number of POTWs affected between about 10 to 20 percent.

There were too few data to adequately evaluate year to year changes in chronic toxicity trigger exceedances for Central Valley POTWs that have a chronic toxicity trigger of greater than 1 TU_c.

An analysis was also conducted to evaluate if seasonality impacted chronic toxicity trigger exceedances. The results of that analysis found that chronic toxicity trigger exceedances did not vary from season to season in the full chronic toxicity testing data set. However, this does not preclude seasonal impacts on individual POTWs.

4.2.1.4 Magnitude of Chronic Toxicity Trigger Exceedances by Percent Effect

Another challenge that arises in evaluating chronic toxicity trigger exceedances using TU_c is the manner (metric) in which they are calculated. As stated previously, chronic toxic units are currently calculated as 100 divided the NOEC, which is one of several statistical measures that may be calculated based on data obtained from the results of dilution series tests (i.e., LC_{50}, EC_{25}, IC_{25}). The NOEC represents the highest effluent concentration of the test at which there are no observed effects on the test organisms. If a dilution series is conducted for the chronic toxicity test with 100 and 50 percent effluent concentrations and toxicity is observed at the 100 percent concentration, but not the 50 percent concentration, the NOEC is 50 percent and the TU_c metric is calculated to be 2 (100 divided by 50). What is not captured in the chronic toxicity unit calculation is the degree to which the test organisms were inhibited in the test where a statistical difference with the control water result was obtained (in this example, at 100 percent concentration). Therefore, the TU_c metric may not reflect the actual effects experienced by test organisms or the potential impact to aquatic life in the receiving waters. The test organisms may have performed relatively well, with small percentage differences between the effluent and the control (i.e., low percent effect), which still yielded a finding of toxicity due to a statistically significant difference between the samples. Alternatively, the test organisms may have performed poorly in the effluent relative to the control (i.e., high percentage difference). Under either of these scenarios, the reported TU_c would still be 2, which is generally an exceedance of the chronic toxicity trigger for most Central Valley POTWs.

Although the TU_c metric has been used to evaluate chronic toxicity of effluent, the State Water Board's 2012 proposed Statewide Toxicity Plan proposed to no longer use the TU_c metric as the benchmark for evaluating chronic toxicity data and assessing compliance with the Basin Plan water quality objectives. The proposed Statewide Toxicity Plan recommends use of a percent "effect" metric to assess the magnitude of chronic toxicity by comparing differences in toxicity test results between the effluent and control water(s).

The 2012 proposed Statewide Toxicity Plan approach for assessing the magnitude of chronic toxicity using the percent effect between the effluent and control was also evaluated using the available test results. Additionally, in the Central Valley Water Board's NPDES General Order (CAG585001, Order No. R5-2017-0085), toxicity is assessed using both TU_c and percent effect.[6]

For the purpose of this evaluation, the percent effect was divided into four groups: less than 25 percent reduction between the effluent and control; between 25 and 50 percent reduction between the effluent and control; between 50 and 75 percent reduction between the effluent and control; and greater than 75 percent reduction between the effluent and control.

---

[6] Central Valley Regional Water Quality Control Board. *Waste Discharge Requirements for Municipal Wastewater Dischargers that Meet Objectives/Criteria at the Point of Discharge to Surface Water* (NPDES No. CAG585001, Order No. R5-2017-0085). August 11, 2017. Available at: https://www.waterboards.ca.gov/centralvalley/board_decisions/adopted_orders/general_orders/r5-2017-0085.pdf, Last accessed October 23, 2018.

A summary of the percent effect on the subset of chronic toxicity data where an exceedance of the chronic toxicity trigger was observed for *C. dubia* reproduction is presented in **Table 5**. When chronic toxicity was indicated for *C. dubia* reproduction (i.e., the chronic toxicity trigger was exceeded), approximately a quarter of those tests showed effects less than 25 percent, and approximately three-quarters of those tests showed effects less than 50 percent.

Table 5. Percent Effect for *Ceriodaphnia dubia* Reproduction Chronic Toxicity Trigger Exceedances for Central Valley POTWs, January 2011 to March 2017.

Control Water	Number of Chronic Toxicity Trigger Exceedances	<25% Reduction	25-50% Reduction	50-75% Reduction	>75% Reduction
Central Valley POTWs with Chronic Toxicity Trigger = 1 TU_c (1)					
Laboratory water (2)	104	31 (29.8%)	48 (46.2%)	13 (12.5%)	11 (10.6%)
Receiving water	27	6 (22.2%)	14 (51.9%)	4 (14.8%)	3 (11.1%)
Central Valley POTWs with Chronic Toxicity Trigger > 1 TU_c (3)					
Laboratory water	4	1 (25.0%)	2 (50.0%)	1 (25.0%)	0 (0.0%)
Receiving water	13	0 (0.0%)	0 (0.0%)	0 (0.0%)	13 (100%)

(1) There are 59 Central Valley POTWs that have a chronic toxicity trigger of 1 TU_c for *C. dubia* reproduction.
(2) There was one toxicity test report that indicated laboratory water was used, but the raw reproduction data were not included, which results in the columns not summing to the total.
(3) There are 7 Central Valley POTWs have a chronic toxicity trigger greater than 1 TU_c for *C. dubia* reproduction.

Similarly, a breakdown of the percent effect on the subset of chronic toxicity data where exceedance of the chronic toxicity trigger was observed for *S. capricornutum* growth is presented in **Table 6**. When laboratory water was used as the control, approximately half of the tests exceeding the chronic toxicity trigger had effects less than 25 percent, with three-quarters of those exceeding the trigger showing effects less than 50 percent. When the receiving water was used as the control, approximately one-third of the tests with chronic toxicity trigger exceedances showed effects less than 25 percent, while approximately two-thirds showed effects less than 50 percent.

POTWs with a chronic toxicity trigger of greater than 1 TU_c generally did not appear to have issues with *S. capricornutum* growth.

Table 6. Percent Effect for *Selenastrum capricornutum* Growth Chronic Toxicity Trigger Exceedances for Central Valley POTWs, January 2011 to March 2017.

Control Water	Number of Chronic Toxicity Trigger Exceedances	<25% Reduction	25-50% Reduction	50-75% Reduction	>75% Reduction
Central Valley POTWs with Chronic Toxicity Trigger = 1 TU_c (1)					
Laboratory water (2)	38	20 (52.6%)	9 (23.7%)	5 (13.2%)	4 (10.5%)
Receiving water	38	13 (34.2%)	11 (28.9%)	11 (28.9%)	3 (7.9%)
Central Valley POTWs with Chronic Toxicity Trigger > 1 TU_c (3)					
Laboratory water	0	–	–	–	–
Receiving water	2	2 (100%)	0 (0.0%)	0 (0.0%)	0 (0.0%)

(1) There are 59 Central Valley POTWs that have a chronic toxicity trigger of 1 TU_c for *S. capricornutum.* Before December 1, 2014, the City of Woodland had a chronic toxicity trigger of 1 TU_c for *S. capricornutum.* Chronic toxicity data collected for the City of Woodland prior to this date were included in this subset of data.

(2) There was one toxicity test report that indicated laboratory water was used, but the raw reproduction data were not included, which results in the columns not summing to the total.

(3) There are 7 Central Valley POTWs that have a chronic toxicity trigger of greater than 1 TU_c for *S. capricornutum*. After December 1, 2014, the City of Woodland has a chronic toxicity trigger of 2 TU_c for *S. capricornutum*. Chronic toxicity data collected for the City of Woodland after this date were included in this subset of data.

It is generally acknowledged that a persistent effect of at least 30 to 50 percent in a sample when compared to the control is desirable for successful TIE implementation (USEPA 1996, USEPA 2007). The results for *C. dubia* where more than one-quarter of the chronic toxicity trigger exceedances observed during the study period would not be anticipated to yield successful TIE outcomes. For *S. capricornutum*, more than one-third of the chronic toxicity trigger exceedances fall in this category. These levels of toxicity may have resulted in POTWs expending extensive resources to investigate low-level toxicity that would not be anticipated to be resolved.

4.2.2 POTW Treatment Level

The level of wastewater treatment provided by POTWs was a factor that was evaluated to determine if it affected chronic toxicity test results. As discussed in **Section 2**, Central Valley POTWs were separated into three groups depending on the level of treatment provided: secondary, advanced secondary, and tertiary treatment plants. Secondary treatment of wastewater typically involves physical treatment through primary clarification and biological treatment through activated sludge, oxidation ditches, and/or trickling filters followed by disinfection and disposal. Advanced secondary treatment includes secondary treatment of wastewater as well as conversion of nitrogen compounds through nitrification and/or denitrification, followed by disinfection. Tertiary treatment involves secondary treatment followed by media or membrane filtration and disinfection and may or may not include nitrification and denitrification. A hypothesis that was tested is whether higher levels of wastewater treatment reduced the frequency of chronic toxicity trigger exceedances. A summary of chronic toxicity trigger exceedances for *C. dubia* reproduction and *S. capricornutum* growth by POTW treatment level is presented in **Table 7**.

Table 7. Chronic Toxicity Trigger Exceedances for Central Valley POTWs by Level of Treatment.

	Secondary	Advanced Secondary	Tertiary
Central Valley POTWs with Chronic Toxicity Trigger = 1 TU_c			
Ceriodaphnia dubia (reproduction)			
Number of Routine Toxicity Tests	92	50	678
Number of Toxicity Trigger Exceedances	23 (25.0%) (1)	9 (18.0%)	99 (14.6%)
Number of POTWs Impacted/Number of POTWs (%)	4/11 (36.4%)	2/5 (40.0%)	32/44 (72.7%)
Selenastrum capricornutum (growth)			
Number of Routine Toxicity Tests	92	48	695 (2)
Number of Toxicity Trigger Exceedances	14 (15.2%)	3 (6.3%)	59 (8.5%)
Number of POTWs Impacted/Number of POTWs (%)	3/11 (27.3%)	1/5 (20.0%)	25/44 (56.8%)
Central Valley POTWs with Chronic Toxicity Trigger > 1 TU_c			
Ceriodaphnia dubia (reproduction)			
Number of Routine Toxicity Tests	99	0	39
Number of Toxicity Trigger Exceedances	14 (14.1%) (3)	–	3 (7.7%)
Number of POTWs Impacted/Number of POTWs (%)	2/3 (66.7%)	–	1/4 (25.0%)
Selenastrum capricornutum (growth)			
Number of Routine Toxicity Tests	104	0	48 (4)
Number of Toxicity Trigger Exceedances	0 (0.0%)	–	2 (4.2%)
Number of POTWs Impacted/Number of POTWs (%)	0/3 (0.0%)	–	1/5 (20.0%)

(1) The City of Davis accounted for 17 of the chronic toxicity trigger exceedances from 33 routine chronic toxicity tests. During the period of the data set evaluated, the City operated a land-based secondary treatment system and underwent facility upgrades. The results from this POTW significantly skew the data set for POTWs with secondary treatment. Excluding the City's data set from this analysis resulted in only 6 chronic toxicity trigger exceedances in 59 routine chronic toxicity tests (10.2 percent).

(2) Before December 1, 2014, the City of Woodland had a chronic toxicity trigger of 1 TU_c for *S. capricornutum*. Chronic toxicity data collected prior to this date were included in this subset of data.

(3) The City of Davis accounted for 3 of the chronic toxicity trigger exceedances from 33 routine chronic toxicity tests. Excluding the City's data set from this analysis resulted in 11 chronic toxicity trigger exceedances in 59 routine chronic toxicity tests (18.6 percent).

(4) After December 1, 2014, the City of Woodland has a chronic toxicity trigger of 2 TU_c for *S. capricornutum*. Chronic toxicity data collected after this date were included in this subset of data.

While higher levels of wastewater treatment also appear to result in fewer chronic toxicity trigger exceedances for *C. dubia* reproduction (Chi-square at $\alpha = 0.5$, p-value = 0.010) for POTWs with a chronic toxicity trigger of 1 TU_c, approximately three-quarters of the exceedances are attributed to the City of Davis, which operated a land-based secondary treatment system and underwent treatment plant upgrades during the period of the data set. Excluding the City's data set from this analysis resulted in POTWs with only secondary treatment having a chronic toxicity trigger exceedance rate of approximately 10.2 percent.

For POTWs with a chronic toxicity trigger of 1 TU_c, higher levels of wastewater treatment appeared to result in fewer chronic toxicity trigger exceedances for *S. capricornutum* growth (Chi-square at $\alpha = 0.5$, p-value = 0.036). Similar to the analysis above, if the City of Davis data set is excluded, POTWs with only secondary treatment had a chronic toxicity trigger exceedance rate of approximately 18.6 percent.

In general, a larger share of higher-level treatment POTWs exhibited chronic toxicity trigger exceedances when compared to lower-level treatment POTWs.

For POTWs that have a chronic toxicity trigger greater than 1 TU_c, the data set is too small to draw any meaningful conclusions as to whether treatment level impacts chronic toxicity testing and the results from the tests.

The level of treatment for Central Valley POTWs with a chronic toxicity trigger of 1 TU_c was also used to evaluate the percent effect observed between the effluent and control when an exceedance of the trigger occurs for *C. dubia* reproduction and *S. capricornutum* growth. A breakdown of the percent effect by POTW treatment level is presented in **Table 8**.

Table 8. Percent Effect for Central Valley POTWs with a Chronic Toxicity Trigger of 1 TU_c by Level of Treatment, January 2011 to March 2017.

Treatment Level	Number of Chronic Toxicity Trigger Exceedances	<25% Reduction	25-50% Reduction	50-75% Reduction	>75% Reduction
Ceriodaphnia dubia (reproduction)					
Secondary	23	6 (26.1%)	12 (52.2%)	2 (8.7%)	3 (13.0%)
Secondary without City of Davis	6	2 (33.3%)	3 (50.0%)	1 (16.7%)	0 (0.0%)
Advanced Secondary	9	4 (44.4%)	5 (55.6%)	0 (0.0%)	0 (0.0%)
Tertiary (1)	99	27 (27.3%)	45 (45.5%)	15 (15.2%)	11 (11.1%)
Selenastrum capricornutum (growth)					
Secondary	14	9 (64.3%)	2 (14.3%)	0 (0.0%)	3 (21.4%)
Secondary without City of Davis	11	7 (63.6%)	1 (9.9%)	0 (0.0%)	3 (27.3%)
Advanced Secondary	3	0 (0.0%)	1 (33.3%)	2 (66.7%)	0 (0.0%)
Tertiary	59	24 (40.7%)	17 (28.8%)	14 (23.7%)	4 (6.8%)

(1) There was one toxicity test report that indicated laboratory water was used, but the raw reproduction data were not included, which results in the columns not summing to the total.

For *C. dubia* reproduction, the level of treatment provided by Central Valley POTWs with a chronic toxicity trigger of 1 TU_c appeared to have little influence on the percent effect observed between the effluent and control. When there was an exceedance of the chronic toxicity trigger, approximately three-quarters of the toxicity tests demonstrated less than 50 percent reduction in the number of offspring produced in effluent tests compared to reproduction in the control for both secondary- and tertiary-level POTWs.

Approximately three-quarters of the chronic toxicity trigger exceedances for *S. capricornutum* growth resulted from less than 50 percent reduction in algal cells in the effluent compared to the control for secondary-level POTWs. Approximately 70 percent of the chronic toxicity trigger exceedances for *S. capricornutum* growth resulted from less than 50 percent reduction in algal cells in the effluent compared to the control for tertiary-level POTWs. This may portend to potential impacts (e.g., potential for biostimulation, more potential for test variability) related to the use of receiving water as the control in the majority of chronic toxicity trigger exceedances observed by tertiary-level POTWs that were previously discussed.

4.2.3 Nitrogen Treatment

Two potential types of treatment processes that may impact chronic toxicity testing for *C. dubia* and *S. capricornutum* are nitrogen treatment (nitrification and denitrification), and disinfection methods. Nitrogen treatment is discussed in this section and disinfection methods are discussed in the following section.

An analysis of nitrogen treatment was conducted specifically because ammonia is a known cause of toxicity. In chronic toxicity testing, samples are analyzed for ammonia prior to test initiation as it can be lethal to the test organisms. The presence of ammonia during testing can also be magnified through pH drift during chronic toxicity testing, which can contribute to artifactual toxicity. If ammonia is present, chronic toxicity testing may be controlled for pH, and in some scenarios, having ammonia removed from the sample prior to testing to remove this interference.[7]

Nitrification is the process of converting ammonia, which is present in domestic wastewater and potentially toxic to aquatic life, into nitrate (and potentially nitrite if there is incomplete nitrification). Nitrate, which is a human health concern particularly to infants, can be converted by denitrification to nitrogen gas, which removes nitrogen from solution to the atmosphere. Since nitrogen treatment occurs in many Central Valley POTWs, the potential impact of these treatment processes were evaluated to determine if they could impact chronic toxicity test results.[8]

After a superficial review of the treatment systems implemented at the 66 POTWs included in this study, it was determined that 18 POTWs do not provide either nitrification or denitrification treatment, 14 POTWs provide only nitrification (i.e., conversion of ammonia to nitrate and/or nitrite), and 34 POTWs provide both nitrification and denitrification. The level and amount of nitrification and/or denitrification provided by a POTW can vary, and may not be entirely represented in these general classifications. Further evaluation may be conducted in the future to refine these classifications. For the purpose of this analysis, the classification of nitrification and denitrification treatment was obtained from the most current NPDES permits for each Central

[7] United States Environmental Protection Agency. *Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms, Fourth Edition*. October 2002.

[8] During the denitrification process, some nitrate will remain in solution and in the final effluent.

Valley POTW. It should be noted that some POTWs may have implemented changes in their nitrogen treatment processes subsequent to their most recent NPDES permits.

A summary of the exceedances of chronic toxicity triggers and number of POTWs impacted based on nitrogen removal is presented in **Table 9**.

In Central Valley POTWs that have a chronic toxicity trigger of 1 TU_c, increased nutrient treatment appeared to reduce the frequency in which exceedances of the chronic toxicity trigger occurred for *S. capricornutum* growth, but may not reduce the frequency in which exceedances of the chronic toxicity trigger for *C. dubia* reproduction being observed. The limited available data and number of POTWs that have a chronic toxicity trigger of greater than 1 TU_c did not allow for a similar analysis.

Table 9. Chronic Toxicity Trigger Exceedances for Central Valley POTWs by Nitrogen Removal, January 2011 to March 2017.

	None	Nitrification Only	Nitrification/ Denitrification
Central Valley POTWs with Chronic Toxicity Trigger = 1 TU_c			
Ceriodaphnia dubia (reproduction)			
Number of Routine Toxicity Tests	112	161	547
Number of Toxicity Trigger Exceedances	25 (22.3%) (1)	25 (15.5%)	81 (14.8%)
Number of POTWs Impacted/Number of POTWs (%)	5/14 (35.7%)	8/13 (61.5%)	25/32 (78.1%)
Selenastrum capricornutum (growth)			
Number of Routine Toxicity Tests	112	158	565
Number of Toxicity Trigger Exceedances	15 (13.4%) (2)	17 (10.8%)	44 (7.8%)
Number of POTWs Impacted/Number of POTWs (%)	4/14 (25.6%)	6/13 (46.2%)	19/32 (59.4%)
Central Valley POTWs with Chronic Toxicity Trigger > 1 TU_c			
Ceriodaphnia dubia (reproduction)			
Number of Routine Toxicity Tests	104	5	29
Number of Toxicity Trigger Exceedances	14 (13.5%)	3 (60.0%)	0 (0.0%)
Number of POTWs Impacted/Number of POTWs (%)	2/4 (50.0%)	1/1 (100%)	0/2 (0.0%)
Selenastrum capricornutum (growth)			
Number of Routine Toxicity Tests	109	5	29
Number of Toxicity Trigger Exceedances	0 (0.0%)	0 (0.0%)	2 (6.9%)
Number of POTWs Impacted/Number of POTWs (%)	0/4 (0.0%)	0/1 (0.0%)	1/2 (50.0%)

(1) The City of Davis accounted for 17 of the chronic toxicity trigger exceedances from 33 routine chronic toxicity tests for POTWs that do not provide nutrient removal. During the period of the data set evaluated, the City operated a land-based secondary treatment system and underwent facility upgrades. The results from this POTW significantly skew the data set for POTWs without nitrogen removal. Excluding the City's data set from this analysis resulted in only 8 chronic toxicity trigger exceedances in 79 routine chronic toxicity tests (10.1 percent).

(2) The City of Davis accounted for 3 of the chronic toxicity trigger exceedances from 33 routine chronic toxicity tests for POTWs that do not provide nutrient removal. Excluding the City's data set from this analysis resulted in 12 chronic toxicity trigger exceedances in 79 routine chronic toxicity tests (15.2 percent).

4.2.4 Disinfection Methodology

As discussed previously, all Central Valley POTWs provide disinfection of treated wastewater either through chlorination or ultraviolet light. A summary of chronic toxicity trigger exceedances based on disinfection methodology and the number of Central Valley POTWs impacted by observed chronic toxicity is presented in **Table 10**.

Table 10. Chronic Toxicity Trigger Exceedances for Central Valley POTWs by Disinfection Methodology, January 2011 to March 2017.

	Chlorination	Ultraviolet
Central Valley POTWs with Chronic Toxicity Trigger = 1 TU_c		
Ceriodaphnia dubia (reproduction)		
Number of Routine Toxicity Tests	260	560
Number of Toxicity Trigger Exceedances	51 (19.6%) (1)	80 (14.3%)
Number of POTWs Impacted/Number of POTWs (%)	12/28 (42.9%)	26/31 (83.9%)
Selenastrum capricornutum (growth)		
Number of Routine Toxicity Tests	263	572
Number of Toxicity Trigger Exceedances	14 (5.3%) (2)	62 (10.8%)
Number of POTWs Impacted/Number of POTWs (%)	6/28 (21.4%)	23/31 (74.2%)
Central Valley POTWs with Chronic Toxicity Trigger > 1 TU_c		
Ceriodaphnia dubia (reproduction)		
Number of Routine Toxicity Tests	104	34
Number of Toxicity Trigger Exceedances	17 (16.3%)	0 (0.0%)
Number of POTWs Impacted/Number of POTWs (%)	3/4 (75.0%)	0/3 (0.0%)
Selenastrum capricornutum (growth)		
Number of Routine Toxicity Tests	109	43
Number of Toxicity Trigger Exceedances	0 (0.0%)	2 (4.7%)
Number of POTWs Impacted/Number of POTWs (%)	0/4 (0.0%)	1/4 (25.0%)

(1) The City of Davis accounted for 17 of the chronic toxicity trigger exceedances from 33 routine chronic toxicity tests for POTWs providing chlorination. During the period of the data set evaluated, the City operated a land-based secondary treatment system and underwent facility upgrades. The results from this POTW significantly skew the data set for POTWs with chlorination. Excluding the City's data set from this analysis resulted in 34 chronic toxicity trigger exceedances in 227 routine chronic toxicity tests (15.0 percent).

(2) The City of Davis accounted for 3 of the chronic toxicity trigger exceedances from 33 routine chronic toxicity tests for POTWs providing chlorination. Excluding the City's data set from this analysis resulted in 11 chronic toxicity trigger exceedances in 230 routine chronic toxicity tests (4.8 percent).

Central Valley POTWs are essentially divided evenly between chlorination and ultraviolet light as the methodology for disinfection (see **Figure 2**). For *C. dubia* reproduction, a higher percentage of chronic toxicity trigger exceedances occurred for POTWs using chlorination compared to POTWs using ultraviolet light. However, if the City of Davis data set is excluded, there does not appear to be a difference in chronic toxicity trigger exceedances between POTWs providing chlorination and ultraviolet light disinfection. It is important to note that nearly every POTW using ultraviolet light disinfection has observed an exceedance of its chronic toxicity trigger for *C. dubia* compared to about half of those POTWs using chlorination.

The results for *S. capricornutum* growth-related toxicity trigger exceedances were even more striking depending on the disinfection methodology. The chronic toxicity trigger was exceeded at twice the rate for ultraviolet light disinfection when compared to chlorination, and approximately three-quarters of the POTWs using ultraviolet light disinfection experienced an exceedance of the chronic toxicity trigger compared to approximately one-fifth of the POTWs using chlorination disinfection. This indicates that the type of disinfection methodology may negatively impact *S. capricornutum* reproduction.

4.3 ACCELERATED TESTING ANALYSIS

As discussed in the **Section 3**, NPDES permits require that POTWs initiate accelerated testing if there is an exceedance of the chronic toxicity trigger. Accelerated testing consists of up to four follow-up chronic toxicity tests within two weeks of each other for the test organism endpoint for which the exceedance occurred. If, after four tests, there is no further indication of toxicity, then the POTW can return to its routine monitoring schedule. However, if any of the chronic toxicity tests conducted during accelerated testing indicates toxicity, and if the POTW has not remedied the problem, the POTW is required to initiate a TRE. In cases where a POTW readily identifies the cause of toxicity during accelerated testing, the NPDES permit allows the POTW to take corrective actions and then re-initiate accelerated testing rather than entering a TRE.

As part of this study, an analysis was conducted to quantify the frequency at which accelerated testing was conducted, how frequently POTWs returned to routine monitoring and how frequently POTWs identified a second chronic toxicity exceedance that required the initiation of a TRE.

In addition to the items discussed previously in this section regarding data sources, in reviewing the chronic toxicity test data from January 2011 to March 2017, the following issues were noted:

- One of the challenges in reviewing accelerated testing data was that POTWs used different nomenclature for how toxicity tests subsequent to routine monitoring were labeled. In some situations, accelerated tests were labeled as TRE- or TIE-related testing or vice versa; and
- Where possible, some POTWs overlapped accelerated testing with routine monitoring. This approach typically helps save costs associated with the extra monitoring, although it introduced some complexity to this data analysis.

Based on the routine chronic toxicity testing summary from **Table 2**, there were approximately 263 instances (i.e., chronic toxicity trigger exceedances) where POTWs may have needed to initiate accelerated testing to follow-up on an indication of toxicity. From the data review and compilation, there were approximately 362 chronic toxicity tests that were characterized as accelerated testing and utilized for this analysis. As discussed previously, accelerated testing reports were not consistently uploaded to CIWQS by POTWs because accelerated testing is a follow-up effort outside of the typically routine information and data that are required to be uploaded to CIWQS. In these cases, accelerated testing reports may have been submitted directly to the POTW's case handler at the Regional Water Boards. Of the accelerated testing reports that were evaluated as part of this study, 21 were associated with *P. promelas*, 282 were associated with *C. dubia*, and 80 were associated with *S. capricornutum*. It should be noted that the total

accelerated test reports do not sum to the total accelerated testing reports reviewed as part of this study because there were instances where a POTW conducted accelerated testing on multiple test organisms at the same time.

From the available accelerated testing data, a breakdown of the type of follow-up conducted by the Central Valley POTWs is presented in **Figure 9**. Approximately 40 percent of the POTWs proceeded with a TRE because there was a second exceedance of the chronic toxicity trigger during accelerated testing. Fourteen percent of the POTWs did not experience a second chronic toxicity trigger exceedance during accelerated testing and returned to routine chronic toxicity testing while three percent were still in the process of conducting accelerated testing at the end of the data period evaluated.

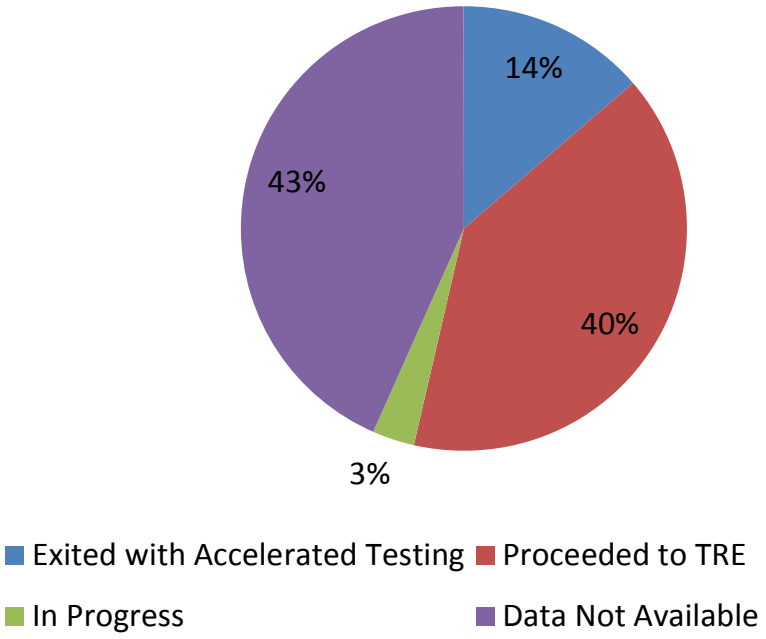

Figure 9. Breakdown of Follow-up Activities to Chronic Toxicity Trigger Exceedances

For 43 percent of the chronic toxicity trigger exceedances, accelerated testing results were not available for evaluation in this study. Some reasons for this may include any or all of the following:

- The toxicity test reports related to follow-up actions were not available through the data sources used for this study. In the data collection phase of this study, some POTWs provided follow-up information in cover letters of required reporting to the Central Valley Water Board without toxicity test attachments;
- There may have been an easily discernible cause of the toxicity that was identified and corrected;
- The discharge may have been seasonal or discontinued and monitoring was not appropriate after the discharge ceased; and/or

- A routine chronic toxicity test sample may have been contaminated or identified as not being representative of the discharge.

To help further refine the analysis of the outcomes of accelerated testing, an additional study would need to focus on this portion of unknown data to better understand how POTWs handled accelerated testing and how many of these POTWs entered the TRE process.

4.4 TOXICITY REDUCTION EVALUATION/TOXICITY IDENTIFICATION EVALUATION ANALYSIS

Information pertaining to completed TREs for Central Valley POTWs was compiled. Unlike obtaining self-monitoring report data from a central repository such as CIWQS, obtaining detailed information on TRE experiences, TRE strategies, and TRE outcomes was challenging, as this information is not submitted in a consistent manner or stored in a centralized location. Information for 39 completed TREs were compiled based on contributions from the special study consultant team and Central Valley Water Board and through a data solicitation directed to CVCWA member agencies. These 39 completed TREs represented the experience of 25 different POTWs; 9 POTWs provided information on more than one TRE.

Four of the 39 TREs were concluded after it was determined that the results of the toxicity testing falsely indicated toxicity. In all four cases, there was a notable test interference or test method protocol deviation identified. Test interferences included a) plating of algae cells (i.e., sticking of algae cells to the walls of the test container) in the *S. capricornutum* test, b) pathogen effects on *P. promelas,* c) biostimulatory receiving water used in *C. dubia* tests, and d) test method protocol deviation for dissolved oxygen in *C. dubia* tests. In each case, evidence of the interference of method deviation was used as the basis to conclude each TRE. Because these four TREs were not related to actual effluent toxicity, they were excluded from the data set that was evaluated.

The test organisms and endpoints subject to the 35 TREs are presented in **Figure 10**. Four of the 35 TREs were associated with POTWs that have a chronic toxicity trigger of greater than 1 TU_c, while the remaining POTWs have a chronic toxicity trigger of 1 TU_c.

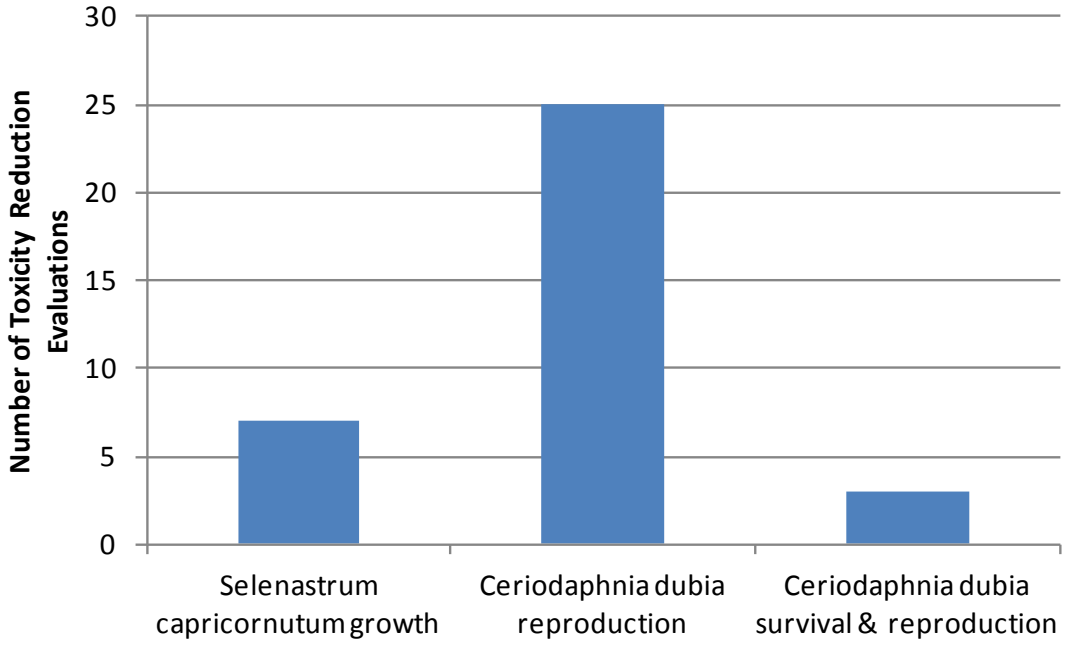

Figure 10. Central Valley POTW Toxicity Reduction Evaluation Test Organism and Endpoint.

From the remaining 35 TREs, generalizations as to the identified cause of toxicity can be made, in addition to generalizations as to which tier of the TRE process (see **Figure 4**) was most informative in advancing the TRE's conclusion. Information was also obtained regarding the duration of the TREs. This evaluation is limited to these generalizations, as each TRE is unique in regard to the POTW affected, the circumstances triggering the TRE, and the level of experience of the parties managing the TRE investigation.

For the 35 completed TREs compiled, identified causes of toxicity were categorized as follows (**Figure 11**):

- POTW maintenance or construction;
- Collection system maintenance or construction;
- Significant industrial user (SIU or other non-domestic or non-commercial user);
- POTW chemical usage;
- An identified contaminant; or
- Unknown cause.

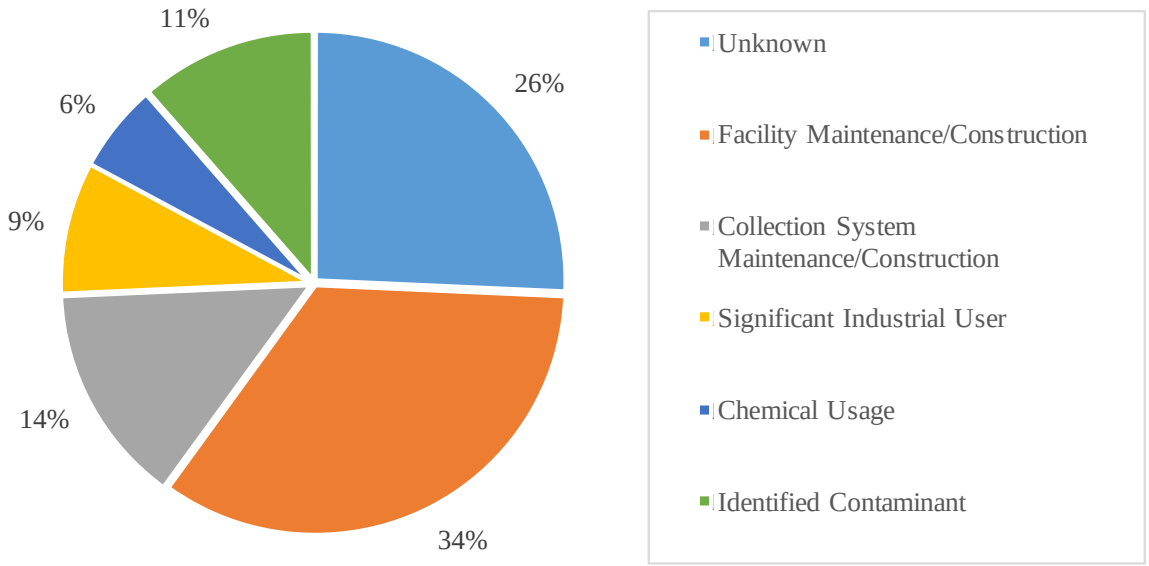

Figure 11. Identified Cause of Toxicity in 35 Toxicity Reduction Evaluations for Central Valley Publicly-Owned Treatment Works.

Of these causes, activities related to POTW or collection system maintenance or construction were identified as the cause of effluent toxicity in 48 percent of the TREs (i.e., 17 of 35). In addition to these cause and POTW chemical usage, about half of the chronic toxicity trigger exceedances may have been potentially controllable. In 26 percent of the TREs (i.e., 9 of 35), no cause for the observed toxicity could be identified. In each of those cases, toxicity was shown to be absent in multiple follow-up chronic toxicity tests. Of the nine TREs where no cause for observed toxicity could be identified, five of these TREs were associated with a relative effect (i.e., difference between the effluent sample and control) of less than 25 percent at the POTWs' IWC in at least one of the two tests required to trigger a TRE. The remaining four of these TREs were associated with a relative effect of 50 percent or greater at the POTWs' IWC. Additional information would be required to further evaluate these individual cases of an "unknown cause" in order to understand the individual drivers behind these inconclusive TREs.

When generalized in terms of the TRE step most informative in advancing the TRE's conclusion (**Figure 12**), the Facility Performance Evaluation (Tier 2) was most influential, and in all but a single case, the accompanying TRE was concluded in less than one year. In contrast, the TIE step was found to be most influential in only 6 percent of TREs despite the TIE step being a component of 34 percent of all TREs. In other words, of 12 TREs that implemented the TIE step, in only 2 cases did the TIE protocol advance the TRE to a conclusion.

Where TIE laboratory reports were available in the compiled CIWQS database, unsuccessful TIEs (i.e., those leading to little or no substantive information) were primarily associated with diminished or absent baseline toxicity. Most TIEs are performed on samples first demonstrated

to be toxic, and thus the TIE is performed on an aged sample (typically more than seven days from sample collection). For labile contaminants subject to degradation or transformation, toxicity can diminish quickly in effluent held in storage, thus rendering the majority of TIE treatments inconclusive, as acknowledged by USEPA (2007):

> *"If the baseline test does not show consistent, measurable toxicity, then one cannot perform a TIE, as the effect of the manipulations on toxicity cannot be assessed."*

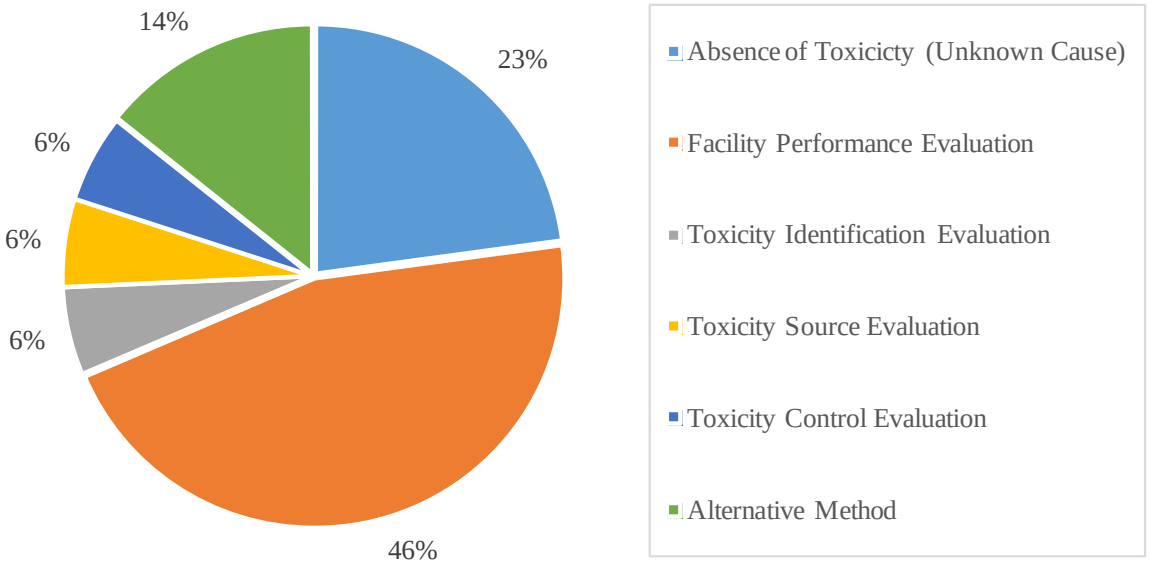

Figure 12. Tier Most Informative as to Successful Conclusion of 35 Toxicity Reduction Evaluations.

While the available TIE laboratory reports were insufficient to determine the threshold of effect at which TIEs are generally successful, it is generally acknowledged by USEPA that a persistent effect of approximately 30 to 50 percent in a sample is desirable for successful TIE implementation (USEPA 1996, USEPA 2007). TIE success can also depend on the experience of the laboratory performing the TIE manipulations and the breadth of manipulations employed.

In 23 percent of the TREs, the cause of toxicity could not be identified before effluent toxicity was no longer observed. Lastly, in 14 percent of TREs, the conclusion was advanced by methods unrelated to any specific tier of the TRE protocol. Of these TREs (5 total cases), the alternative methods included:

- Instream bioassessment (2 cases);
- Comparative split laboratory testing (1 case); and
- Special request of regulatory authority to transition to a Toxicity Evaluation Study (2 cases).

As shown in **Figure 13**, of the 35 TREs evaluated, 77 percent were concluded within 1 year of initiation, with 9 percent of TREs exceeding 3 years. Of the three TREs that exceeded three years, two were concluded after substantial infrastructure improvements were constructed, including installation of new treatment processes. In these two cases, the duration of the TRE is partially explained by time involved in the design, financing, and/or construction of new or upgraded treatment facilities.

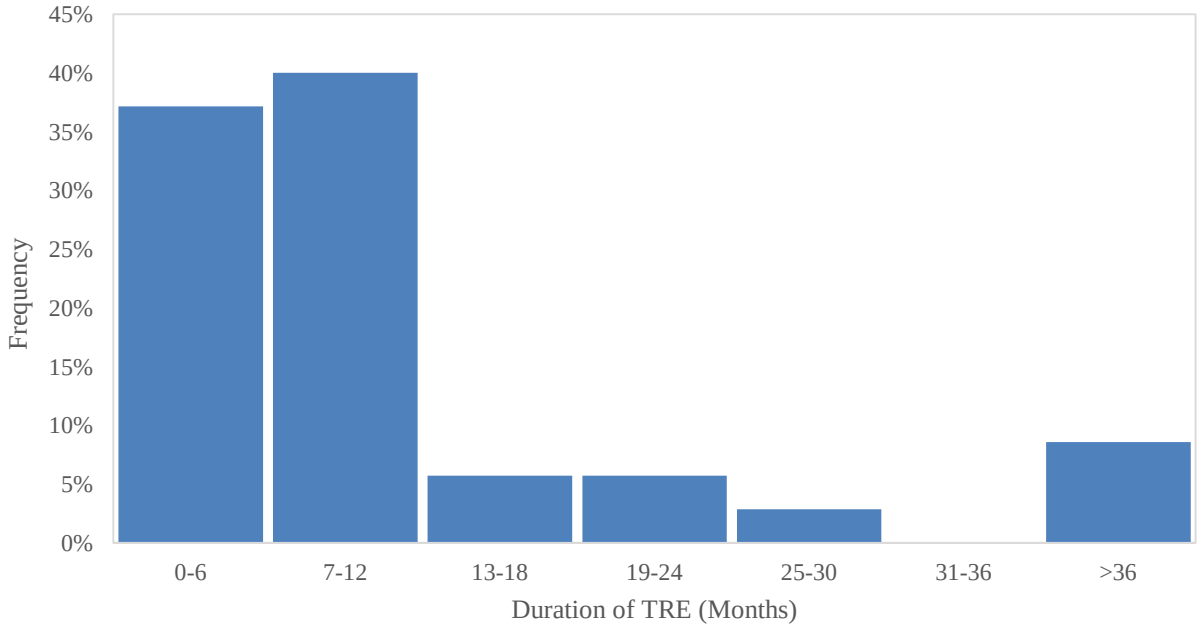

Figure 13. Duration of Toxicity Reduction Evaluations.

4.4.1 TRE Indicated Causes for Effluent Toxicity at Municipal POTWs

The Facility Performance Evaluation is an essential component of the TRE protocol. As shown in **Figure 12**, the Facility Performance Evaluation can itself frequently advance a TRE to conclusion, or can facilitate an efficient focusing of subsequent TRE efforts such as TIE and Toxicity Source Evaluation. A summary of potential causes of effluent toxicity at a POTW, as derived from the completed TREs compiled for this assessment, is provided in **Table 11**. These potential causes should be considered when performing a Facility Performance Evaluation.

Table 11. TRE Indicated Causes for Toxicity.

Treatment Facility Operations, Maintenance, or Construction
1. Identified upset or treatment process malfunction 2. Non-routine maintenance or repair activities - equipment failure - refurbishment or replacement of equipment - repurposing or reuse of idled or unused equipment or facilities - removal/abatement of nuisance conditions (biofilms, struvite, algae blooms, invasive organisms) - accidental or emergency bypass 3. Routine maintenance activities - filter media replacement - filter media backwash - intentional bypass - equipment replacement and preventative maintenance 4. Changes in treatment process standard operating procedures 5. Construction activities - intentional bypass - stress testing - treatment process start-up or optimization 6. Extreme weather events
Treatment Facility Chemical Usage
1. Treatment chemical overdose 2. Treatment chemical spill 3. Treatment chemical optimization or chemical alternative trials
Collection System Operations, Maintenance, or Construction
1. Root control (mechanical and chemical) 2. Hydrocleaning 3. Trenchless repair/slip lining 4. Equipment failure and repair (influent flow spiking due to lift station failure and repair) 5. Extreme weather events
Significant Industrial Users and Pretreatment
1. Changes in industrial user base (addition or removal of industrial user) 2. Modified operations of and/or discharge from industrial user 3. Malfunction or insufficient pretreatment
Other
1. Test interferences (low hardness effluent and pathogens) 2. Identified chemical contaminants (metals, ammonia) 3. Inter-laboratory variability

4.5 SPLIT-LABORATORY TESTING AND VARIABILITY

Due to an indication of effluent toxicity, a number of POTWs have performed 'split-laboratory' comparisons in which samples are split between two or more laboratories. The primary goal of such evaluations is to determine if the toxicity identified by one laboratory is reproducible by another laboratory testing the same samples. For this study, 'split-laboratory' chronic toxicity testing data were obtained from both Central Valley POTWs and the Sanitation Districts of Los Angeles County (LACSD). For the Central Valley POTWs, a total of four California laboratories were involved in the 'split-laboratory' studies; two laboratories were involved with the LACSD testing (i.e., LACSD laboratory and an unidentified laboratory). So as to maintain the confidentiality of the laboratories, the results were anonymized in this report. All laboratory reports were evaluated using the report review checklist developed for CVCWA as part of this study. The report review checklists are provided in **Appendix A**. It is important to note that the LACSD results could not be fully vetted using the report review checklist because only the statistical analysis printouts were provided. However, the chronic toxicity test results were evaluated to confirm that test acceptability criteria were met for each test.

Chronic toxicity test results were excluded from comparisons between/among laboratories if:

- The chronic toxicity tests did not meet the test acceptability;
- A laboratory deviated from method requirements (e.g., incorrect organism age used, inappropriate substitution of a different control treatment);
- The reference toxicant test was out of range;
- The concurrent reference toxicant did not meet test acceptability criteria; and/or
- The concurrent culture water control did not meet test acceptability criteria.

In addition, laboratories have observed pathogen interferences during performing 'split laboratory' testing for the *C. dubia* test and *P. promelas* test. Pathogen interference in the *P. promelas* test is primarily identified visually via the presence of bacterial/fungal growth on the fish, but may also include the observation of sporadic mortalities that may result in high inter-replicate variability and an elevated coefficient of variation (CV) for the affected treatment. In this split-laboratory evaluation, pathogen interferences for the *C. dubia* tests were not due to visible signs of pathogens present on the test organisms, but rather was inferred based on weight of evidence (i.e., presence of a flat concentration-response relationship with similar effects at all concentrations, and the elimination of other factors that could cause such a response curve). Of *C. dubia* split tests that were not qualified for other reasons, three split tests exhibited a flat concentration-response relationship at one laboratory, but not the other. This resulted in disagreement in the NOEC/IC_{25} determinations. As the laboratory that identified such cases considered the pathogen response to be a test interference, and given the lack of agreement between the laboratories due to this confounding factor, they were excluded from this split-laboratory evaluation. Although other observations of differences between laboratories (e.g., sample arrived at the laboratory the day following sample collection, no reference toxicant test reported) were identified using the report review checklist, these differences did not result in a test being rejected, so such results were further evaluated to determine how frequently the test findings were in agreement. In all, ten split test comparisons were omitted from the *C. dubia* data set and three split tests were omitted from the *P. promelas* data set for one or more of the reasons stated above.

USEPA 2002 requires that laboratories evaluate data for outliers, and to present data both including and excluding the outliers. In reviewing the laboratory reports, it was determined that only one laboratory performed statistical outlier evaluations and reported data including and excluding outliers when they occur. Given that the exclusion of outliers can result in improved test precision and can result in altering differences between laboratories that are compared for 'split-laboratory' testing, the 'split-laboratory' comparison was also performed using the results excluding outliers from the data set.

Sub-lethal endpoint data (e.g., growth, reproduction) were evaluated for the following 'split-laboratory' studies to determine the frequency of consistent findings between/among laboratories:

- *S. capricornutum*: 14 'split-laboratory' studies performed;
- *C. dubia*: 30 'split-laboratory' studies performed; and
- *P. promelas*: 4 'split-laboratory' studies performed.

Two approaches were used to determine how often laboratories agreed during 'split-laboratory' testing. The goal of the first evaluation was to determine if the laboratories agreed whether a POTW was in or out of compliance for the toxicity testing (i.e., exceeded the chronic toxicity trigger or not), specifically by evaluating the toxicity evaluations at the IWC. The goal of the second evaluation was to determine the degree to which laboratories differed in the evaluation of key statistical analyses (e.g., NOEC/EC_{25}).

One of the challenges of the 'split-laboratory' data set used in this evaluation was that the vast majority of the comparisons were limited to only a pair-wise comparison of the results from two laboratories, and there were a limited number of laboratories involved. Although there may be a variety of reasons that only two laboratories may be included in a 'split-laboratory' evaluation (e.g., costs, equipment available for collecting a sufficient volume of effluent), a more powerful conclusion may be drawn by including a greater number of laboratories when possible. Warren Hicks et al., (2000) indicated that comparability among laboratories should be addressed by using ***several*** laboratories (emphasis added, suggesting more than two laboratories) before making a determination on whether an effluent is in compliance. Without additional supporting lines of evidence, a weakness of 'split-laboratory' testing with only two laboratories is the lack of a means for placing a higher value on one set of the test results over another. However, when only two laboratories are used in 'split-laboratory' testing and both laboratories generate acceptable test data but are in disagreement, other information (e.g., other aspects of the TRE, instream bioassessment, ambient toxicity testing) may need to be considered to decide which split-test result is best supported.

4.5.1 Laboratory Agreement – Compliance Determination

The NOEC, which is the statistical compliance metric for Central Valley POTWs, was determined from the laboratory reports and compared against a TU_c calculation to determine if the chronic toxicity trigger was/was not met. In addition, a similar evaluation was performed using the IC_{25} (when it could be calculated), since this is another valid statistical method and, in fact, is the recommended statistical analysis in the USEPA method (USEPA 2002).

For the comparison of toxicity test results between laboratories, it is expected that laboratories will generally agree when effects are low (i.e., less than 25 percent reduction in sub-lethal endpoint) and when effects are high (i.e., greater than 75 percent reduction in sub-lethal endpoint). So as to evaluate if there was a magnitude of effect necessary for laboratories to generally agree on toxicity, 'split laboratory' results were evaluated by grouping the percent effect into four categories (i.e., 25 percent or less effect, 26 to 50 percent effect, 51 to 75 percent effect, and 76 to 100 percent effect) and determining the frequency of compliance agreement between laboratories using both the NOEC and IC_{25} test statistics. For this evaluation of magnitude of effect for laboratories to generally agree on the presence of toxicity, there was only sufficient 'split laboratory' data for the *S. capricornutum* and *C. dubia* tests.

Selenastrum capricornutum 'Split-Laboratory' Results

For the 14 cases examined, the results of the 'split-laboratory' testing indicated that there was 65 percent agreement between the laboratories using the NOEC (**Figure 14**) and 75 percent agreement using the IC_{25} (**Figure 15**) for the *S. capricornutum* test method.

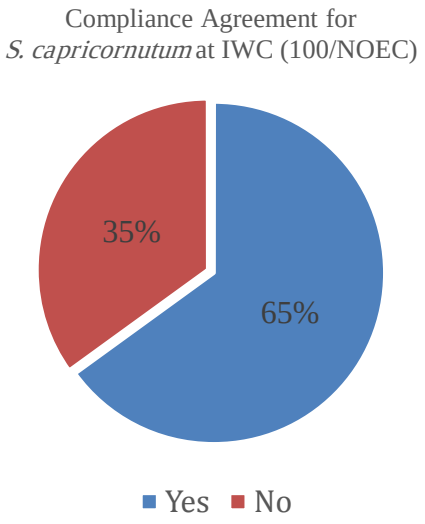

Figure 14. Percent Agreement at IWC (NOEC) for *Selenastrum capricornutum*.

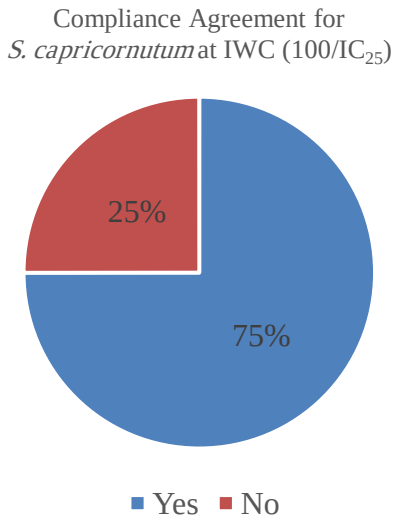

Figure 15. Percent Agreement at IWC (IC_{25}) for *Selenastrum capricornutum*.

As shown in **Figures 16** and **17**, there was a 60 and 67 percent agreement for the NOEC and IC_{25}, respectively, when the laboratories observed a 25 percent reduction or less in algal growth and a 100 percent agreement for both the NOEC and IC_{25} when there was 76 percent or greater effect in algal growth. Although there is no metric where laboratories must agree, the percent agreement when there as a 25 percent or less reduction in algal growth shows only a moderate level of agreement for 'split-laboratory' test results at this percent reduction (i.e., low percent effects).

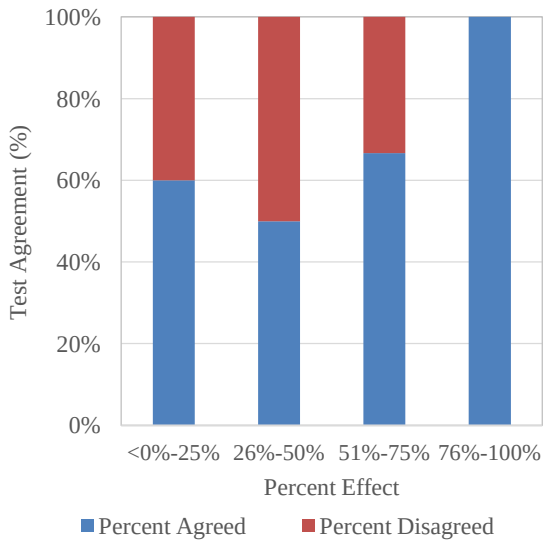

Figure 16. Percent Agreement at IWC (NOEC) based on Percent Effect for *Selenastrum capricornutum*.

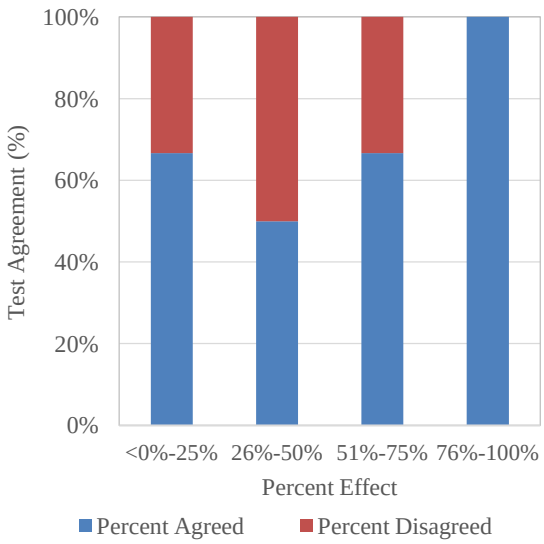

Figure 17. Percent Agreement at IWC (IC_{25}) based on Percent Effect for *Selenastrum capricornutum*.

Ceriodaphnia dubia 'Split-Laboratory' Results

For the 30 cases examined, the results of the 'split-laboratory' testing indicate that there was 73.3 percent agreement between the laboratories using the NOEC (**Figure 18**) and 82.7 percent agreement using the IC_{25} (**Figure 19**) for the *C. dubia* test method; it should be noted that the IC_{25} could not be calculated for 1 of the 30 comparisons. Of the 30 tests that were compared, an outlier was reported in the data set for eight comparisons. If outliers were excluded, the results of the 'split-laboratory' testing indicated that there was 76.7 percent agreement between the laboratories using the NOEC and 79.3 percent agreement using the IC_{25}. The presence/absence of outliers for the *C. dubia* tests therefore had a minor impact on the comparability of 'split-laboratory' testing.

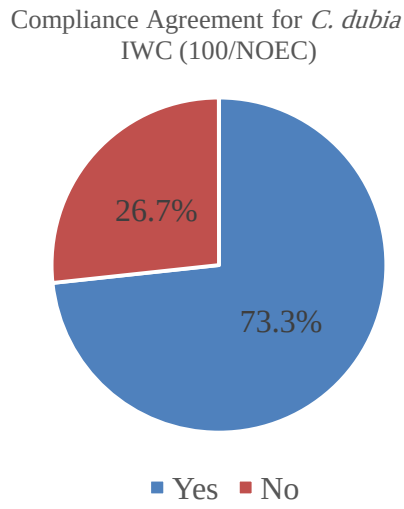

Figure 18. Percent Agreement at IWC (NOEC) for *Ceriodaphnia dubia*.

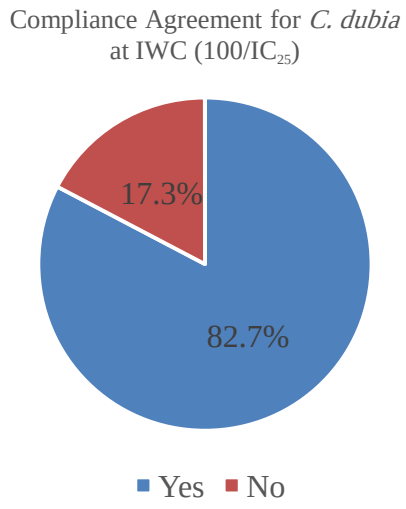

Figure 19. Percent Agreement at IWC (IC_{25}) for *Ceriodaphnia dubia*.

As shown in **Figures 20** and **21**, there was a 78 and 87 percent agreement for the NOEC and IC_{25}, respectively, when the laboratories observed a 25 percent or less reduction in *C. dubia* reproduction and 83 and 100 percent agreement, respectively, for both the NOEC and IC_{25} when there was a 76 percent or more effect in *C. dubia* reproduction. These results show a relatively high level of agreement between laboratories when there was a low and high percent effect on *C. dubia* reproduction.

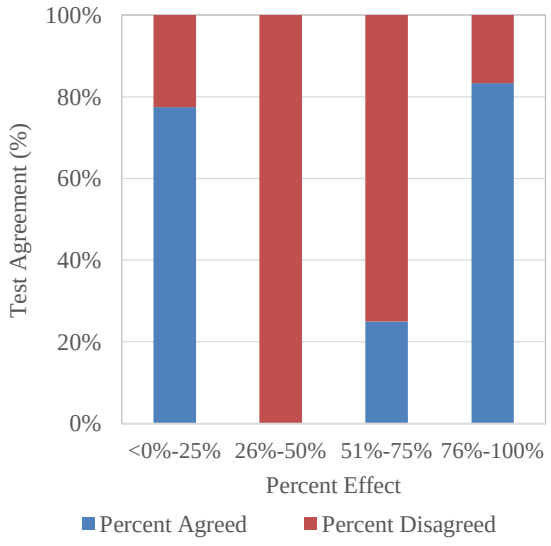

Figure 20. Percent Agreement at IWC (NOEC) based on Percent Effect for *Ceriodaphnia dubia* Reproduction.

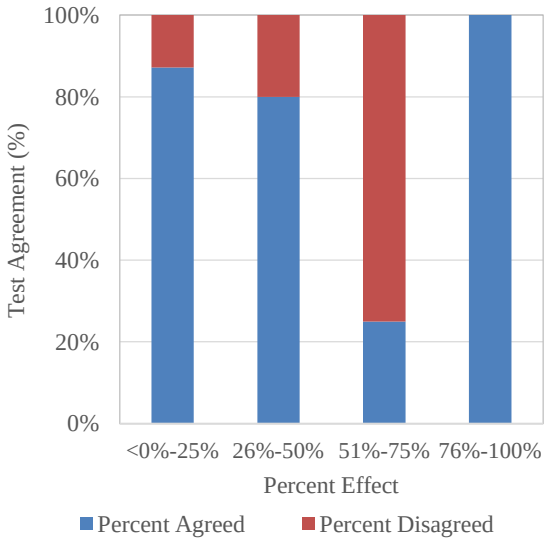

Figure 21. Percent Agreement at IWC (IC_{25}) based on Percent Effect for *Ceriodaphnia dubia* Reproduction

***Pimephales promelas* 'Split-Laboratory' Results**

For the four cases examined, the results of the 'split-laboratory' testing indicate that there was 100 percent agreement between the laboratories using both the NOEC (**Figure 22**) and IC_{25} (**Figure 23**) for the *P. promelas* test method.

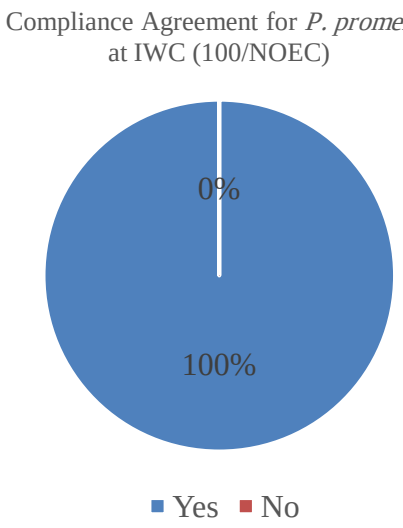

Figure 22. Percent Agreement at IWC (NOEC) for *Pimephales promelas*.

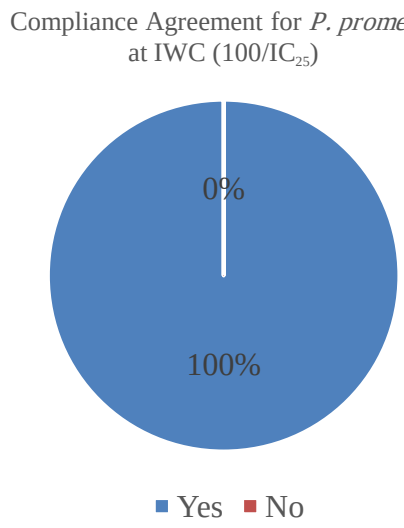

Figure 23. Percent Agreement at IWC (IC_{25}) for *Pimephales promelas*.

4.5.2 Laboratory Agreement – Variability in the NOEC/EC_{25}

Although comparing 'split-laboratory' data for agreement with effluent limitations is a critical evaluation, this approach alone does not provide a metric for the magnitude of differences among laboratory results. For this study, an evaluation of the magnitude of differences was also conducted by comparing the relative percent difference between test results for the EC_{25} and the difference in the number of dilution treatments for the NOEC. The relative percent difference was calculated as follows:

$$RPD = \text{absolute value } (EC_{25} \text{ Lab 1} - EC_{25} \text{ Lab 2}) / \text{average}(EC_{25} \text{ Lab 1}, EC_{25} \text{ Lab 2}) * 100$$

The results for this set of comparisons are provided below.

***Selenastrum capricornutum* 'Split-Laboratory' Results**

The comparison of the relative percent difference between EC_{25} values is provided in **Figure 24**. Test EC_{25} values were within a relative agreement of 25 percent for 89 percent of the tests, while 100 percent of the tests were within a relative agreement 50 percent. For 'split-laboratory' evaluation of the NOEC, the laboratories agreed for 50 percent of the tests, and the NOEC values were within one test treatment for an additional 39 percent of the tests. (**Figure 25**). The number of treatments different reflects the difference in the NOEC based on the dilution between the two laboratories. For example, a difference of 1 means that one laboratory found the NOEC to be 100 percent while the second laboratory found the NOEC to be 75 percent. Similarly, a difference of

2 means that one laboratory found the NOEC to be 100 percent while the second laboratory found the NOEC to be 50 percent.

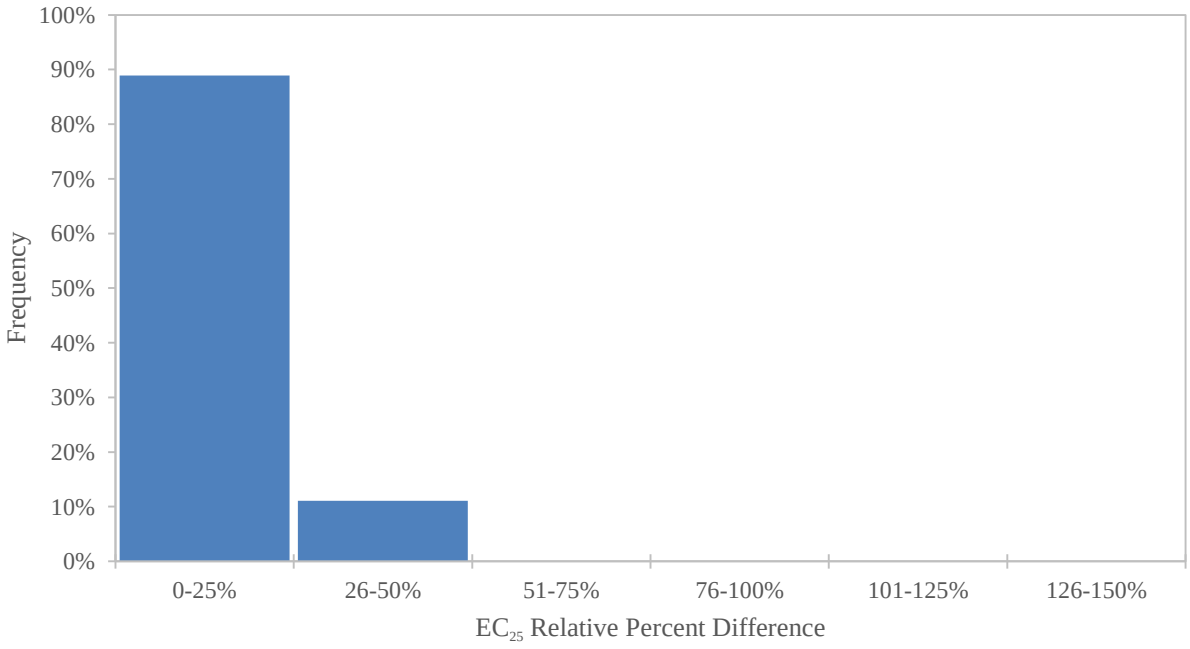

Figure 24. Relative Percent Difference in EC_{25} Values for *Selenastrum capricornutum* 'Split-Laboratory' Tests

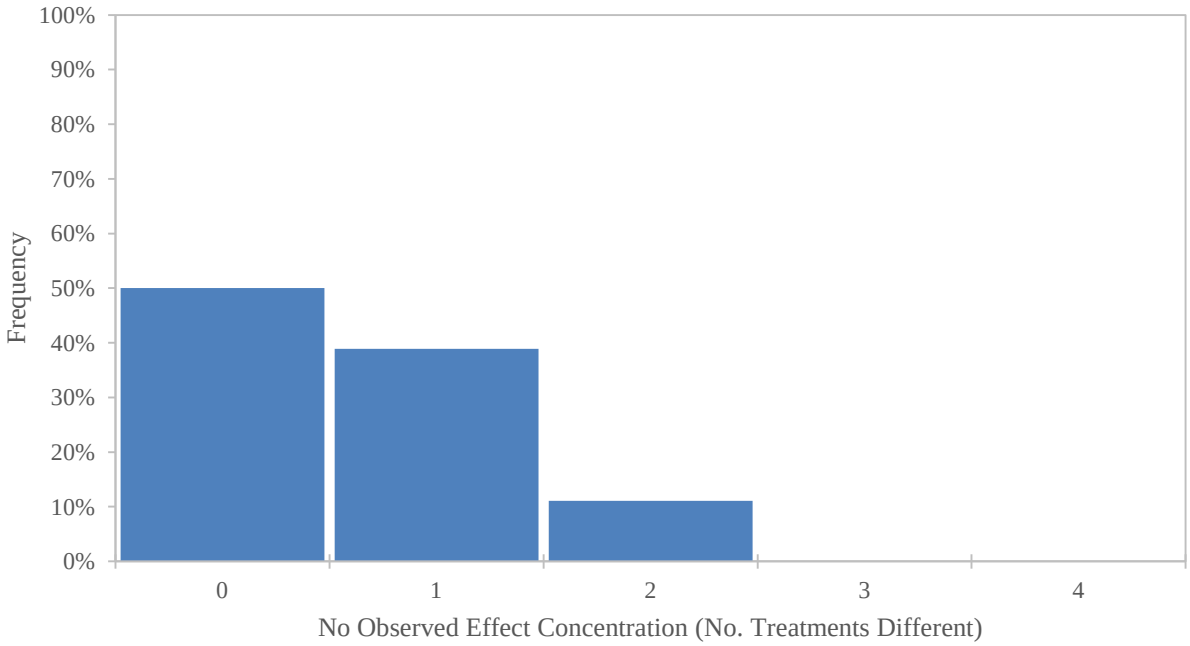

Figure 25. Comparison of NOEC Values for *Selenastrum capricornutum* 'Split-Laboratory' Tests.

Ceriodaphnia dubia 'Split-Laboratory' Results

The comparison of the relative percent difference between EC_{25} values is provided in **Figure 26**. Test EC_{25} values were within a relative agreement of 25 percent for 66 percent of the tests, while 87 percent of the tests were within a relative agreement of 50 percent. For 'split-laboratory' evaluation of the NOEC, the laboratories agreed for 68 percent of the tests, and the NOEC values were within one test treatment for an additional 25 percent of the tests (**Figure 27**).

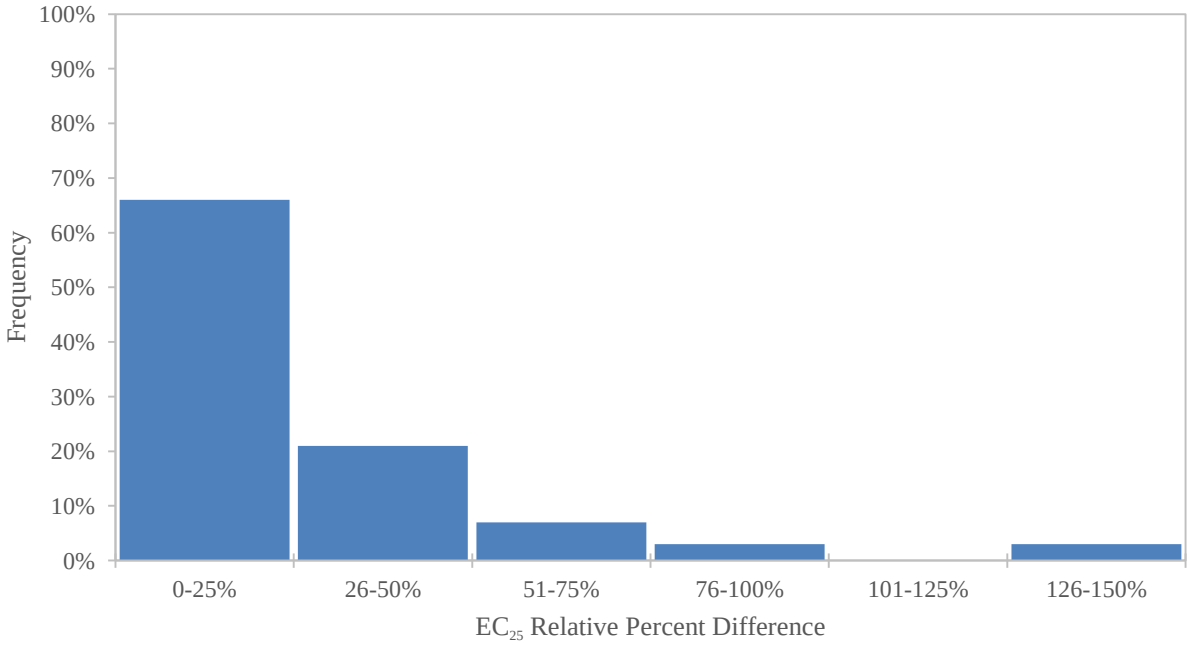

Figure 26. Relative Percent Difference in EC_{25} Values for *Ceriodaphnia dubia* 'Split-Laboratory' Tests.

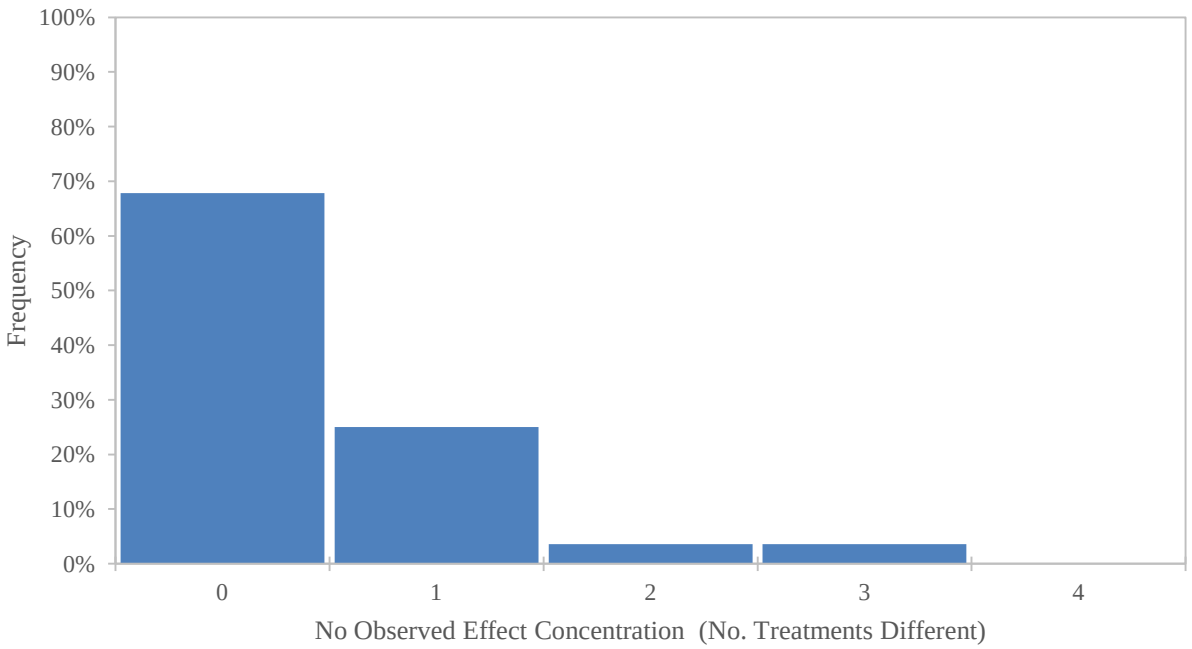

Figure 27. Comparison of NOEC Values for *Ceriodaphnia dubia* 'Split-Laboratory' Tests.

Pimephales promelas 'Split-Laboratory' Results

The comparison of the relative percent difference between EC_{25} values is provided in **Figure 28**. Test EC_{25} values were within a relative agreement of 25 percent for all of the tests. For 'split-laboratory' evaluation of the NOEC, the laboratories agreed for all of the tests (**Figure 29**).

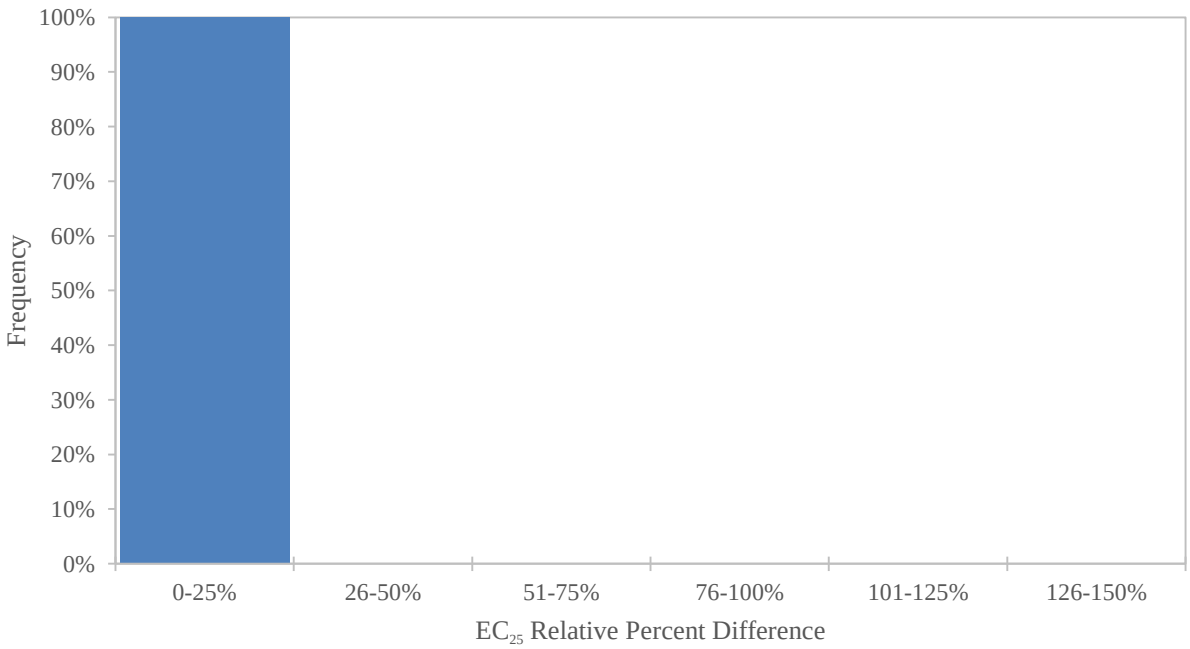

Figure 28. Relative Percent Difference in EC_{25} for *Pimephales promelas* 'Split-Laboratory' Tests.

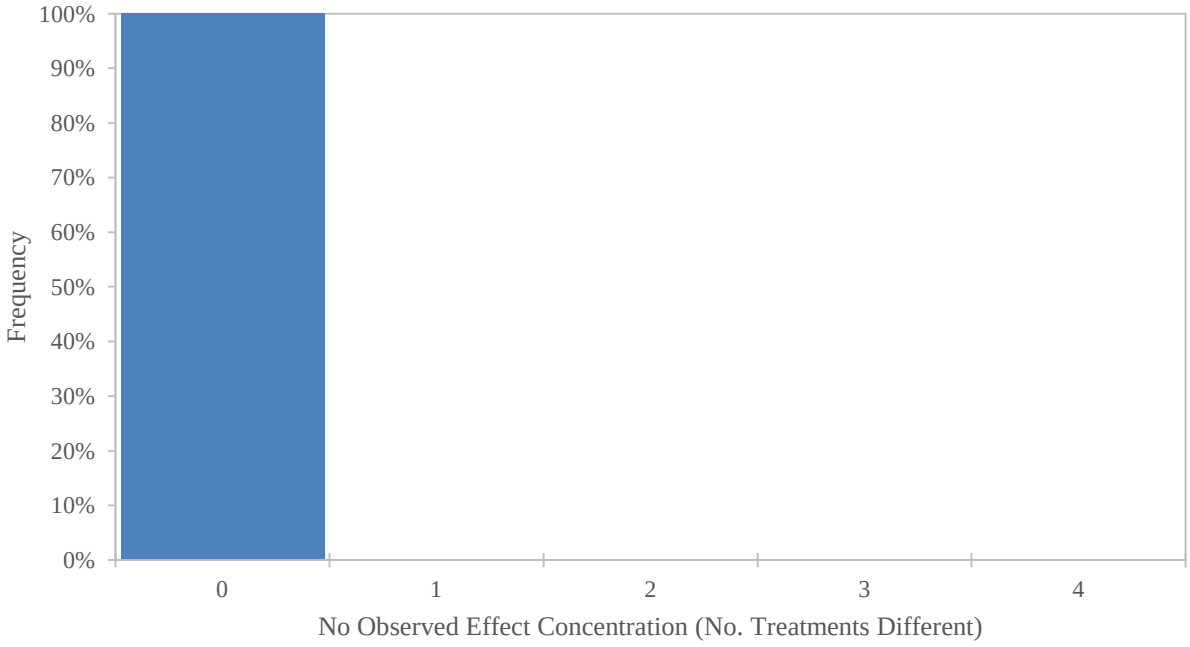

Figure 29. Comparison of NOEC Values for *Pimephales promelas* 'Split-Laboratory' Tests.

4.5.3 Split-Laboratory Testing and Variability Summary

Overall, "split-laboratory" studies resulted in a moderate to high degree of agreement. For the "split-laboratory" comparisons that were performed, the greatest agreement between laboratories occurred for the *P. promelas* test. The laboratories always agreed for this test, but it is important to note that the sample size (n=4) was quite small for this protocol. The *C. dubia* test had the next highest agreement between (and among) laboratories (73-83%), and the lowest agreement between laboratories occurred with the *S. capricornutum* test (65-77%). Typically, there was a slightly greater agreement between laboratories using a comparison of the IC_{25} as when compared to the NOEC.

When only two laboratories were used in "split-laboratory" testing and the laboratories both generate acceptable test data, but are in disagreement, it is unclear which laboratory should be used to assess compliance. A principal limitation of pair-wise "split-laboratory" testing is the lack of a consistent and defensible means of placing higher value on one test results over another when the test results are not in agreement. Development a consistent and defensible approach to "split-laboratory" testing, if "split-laboratory" testing is to be employed with rigor, is recommended. Warren Hicks et al., (2000) recommendations should be considered in which they indicated that comparability among laboratories should be addressed by using ***several*** laboratories (emphasis added) before making a determination on whether an effluent is in compliance.

This page is intentionally left blank.

Section 5. Variability in Sub-Lethal Endpoints

In the previous section, an analysis was conducted to characterize chronic toxicity test results that were observed in discharges from Central Valley POTWs. The characterization identified the frequency of toxicity trigger exceedances and evaluated several potential factors that could influence or impact the outcome of the chronic toxicity test. In this section, an analysis was conducted to identify and evaluate specific variables during chronic toxicity testing that can affect the outcome of a chronic toxicity test.

5.1 VARIABILITY IN CHRONIC TOXICITY TESTING METHODOLOGY

There are three sources of method variability: intra-test, intra-laboratory, and inter-laboratory. Intra-test variability is the variability of the test organism response within a single test. Intra-laboratory variability is the variability of tests conducted over time within the same laboratory, which is affected by intra-test variability. Inter-laboratory variability is the variability among laboratories, which is measured by evaluating the results of different laboratories testing the same sample(s) using the same test method. Inter-laboratory variability is affected by intra-test variability and intra-laboratory variability. The factors that affect these three forms of test variability are provided in **Table 12**. Changes in any of the factors listed in **Table 12** can affect variability (e.g., such as changes in culture practices).

Table 12. Sources of Variability for Aquatic Toxicity Testing Methods.

Category of Variability	Sources of Variability
Intra-test variability	- Replicates: number of replicates and number of organisms per replicate - Culture quality: genetic variability, culture condition - Microbial interferences*:* epibionts (e.g., peritrichs, bacteria.)
Intra-laboratory variability	- Test conditions: selection and variability in food, control/dilution water quality/consistency, consistency of test conditions - Organism condition: culture quality (affected by conditions above) - Laboratory experience: testing facility, quality assurance/quality control (QA/QC) program, analyst training program, variability among analysts - Analyst experience: training support, adherence to protocol
Inter-laboratory variability	- Intra-test variability factors: see above - Intra-laboratory factors: see above - Differences allowed in method: source and type of food, control/dilution water, organism culture condition

Although the chronic toxicity test methods that are in NPDES permits for Central Valley POTWs must be performed by all laboratories following the current testing method manual, each test method provides considerable flexibility for the laboratory in terms of how testing is performed. Factors that may affect test quality across all test methods include the type of control/dilution water (i.e., must support adequate test organism performance, consistently meets minimum test acceptability criteria, is of consistent quality), source of test organisms (i.e., in-house or vendor source), test organism quality, food quality, and culture conditions. Some species-specific

examples that can affect inter-laboratory variability due to the flexibility laboratories have to implement the methods are as follows (USEPA recommendations included parenthetically):

- *Pimephales promelas*
    - Source of food (newly hatched *Artemia salina*)
        - Multiple "grades" of *A. salina* are available
    - Feeding regime (sufficient nauplii are provided in excess)
- *Ceriodaphnia dubia*
    - Feeding regime (0.1 mL each of Yeast Cerophyll®-trout [YCT] and algal suspension per test chamber daily)
    - Food type (YCT and *S. capricornutum*)
        - other foods may be used
    - Culture quality (laboratories must culture species in-house)
        - cultures producing low numbers of offspring or males may not meet test acceptability criteria.
    - Control/dilution water (algal stock culture medium, enriched uncontaminated source of receiving water, synthetic water, dilute mineral water [DMW])
- *Selenastrum capricornutum*
    - Light intensity (400 ± 40 ft-candles) and quality ("cool white" fluorescent)
    - Initial cell density (10,000 cells/ml)
    - Culture quality (labs must culture species in-house)
    - Control/dilution water (algal stock culture medium, enriched uncontaminated source of receiving water, synthetic water, DMW)

The most recent USEPA freshwater chronic toxicity testing manual (Section 4.13.1, EPA 2002) identifies the following sources of variability in toxicity test results:

1. Experience and skill of the laboratory analyst;
2. Test organism age, condition, and sensitivity;
3. Dilution water quality;
4. Temperature control; and
5. The quality and quantity of food provided.

Further, the test method manual indicates that results may depend on the species (i.e., algae more sensitive to herbicides than invertebrates) and the strain and source of organisms.

Warren-Hicks et al., (1999), in a study comparing intra- and inter-laboratory results of test variation, concluded that some laboratories could consistently reproduce test results, while others could not, and inferred that test precision is a factor of laboratory experience. In addition to identifying that laboratory experience is a driver for consistent test results among laboratories, several other studies also identified laboratory and analysts experience/training (emphasis expressed in bold below) as key factors affecting inter-laboratory variability:

- Grothe et al., (1996) indicated the following regarding variability in WET test methods: *"A number of problems with WET tests are caused by misapplication of the tests, misinterpretation of the data,* ***lack of competence of the laboratories conducting WET***

testing, poor condition/ health of test organisms, and lack of training of laboratory personnel….. In addition, an effective QA/QC program will improve data quality and reduce test variability."

- DeGraeve et al., (1998) in a Water Environment Research Foundation- (WERF) sponsored study concluded that *"The project team believes that the results demonstrate that the test methods can be routinely completed successfully by well-trained, competent WET testing laboratories and that the test results, considered collectively, suggest that the test methods that are being used to measure WET are technically sound."*
- Both Grothe et al. (1996) and DeGraeve et al., (1998) also noted that method performance occurs when the methods are closely followed by experienced analysts. USEPA 2000 further states that "*inexperienced individuals can perform analyses incorrectly or fail to follow appropriate methods and quality assurance practices"* and further that *"as with any other analytical system, lack of experience in performing the analyses, adherence to prescribed QA practices, or good laboratory practices will reduce the precision of the tests."*

Although much of the literature identified laboratory and technician experience as key to reducing method variability, and in fact experience and skill of the laboratory analyst was the first item noted in the USEPA method manual, laboratories of comparable experience with highly-trained technicians may still have elevated test variability based on factors noted in **Table 12**, including control water selection, culture condition, and food quality. Although it is hoped that future research will definitively evaluate these parameters as sources of laboratory variability, the following two case studies identify control water selection as a source of test variability.

5.1.1 Case Study 1 – *Selenastrum capricornutum* Control Medium Effect on Toxicity

The USEPA method permits the use of a variety of laboratory control waters for this test, and recommends algal culture control medium, enriched uncontaminated receiving water, synthetic water (i.e., distilled/deionized/Type I water amended with reagent grade salts), or DMW (USEPA 2002). The control water must simply be of consistent quality, meet test acceptability criteria, and be free from contamination. For testing, each treatment, including the control water, is to be amended with nutrients to support algal growth; the nutrients are consistent with those that are used to create the algal culture control medium. This means that any control water selected by a laboratory other than the algal culture control medium may have additional minerals that can serve as an additional nutrient source to *S. capricornutum*. Section 14.3.5 of the method manual indicates that nutrients in the dilution (or control water) may confound test results.

Clark et al. (2011) performed a study in which they evaluated the performance of a variety of acceptable control waters commonly used by laboratories compared to a POTW effluent. The following control waters were tested:

- algal culture control medium (Type I water, with USEPA recommended macro- and micro-nutrients);
- USEPA moderately hard water (EPAMH),
- dilute mineral water (DMW, diluted Perrier water to be moderately-hard),
- dilute mineral water (DMW, diluted Arrowhead water to be moderately-hard),
- dilute mineral water (DMW, 26% Evian:74% Arrowhead to be moderately-hard).

Each water was amended with algal growth nutrients per USEPA method, and consistent with those added to Type I water to make the algal culture control medium.

When the growth of the five acceptable control waters were compared to the effluent treatment, it was determined that the EPAMH and DMW waters were all stimulatory compared to the algal culture medium (per values in the parentheses), and that only the algal culture medium sample identified the effluent as not toxic; all other control waters identified the effluent as toxic (**Figure 30**).

Through an analysis of total macro-and micro-nutrients, the EPAMH and DMW control waters had increased concentrations of bicarbonate, calcium, magnesium, and sodium when compared to the algal culture medium (**Table 13**). Other macro- and micro-nutrients that had increased concentrations for some of the control waters when compared to the algal culture medium were chloride, potassium, and sulfate.

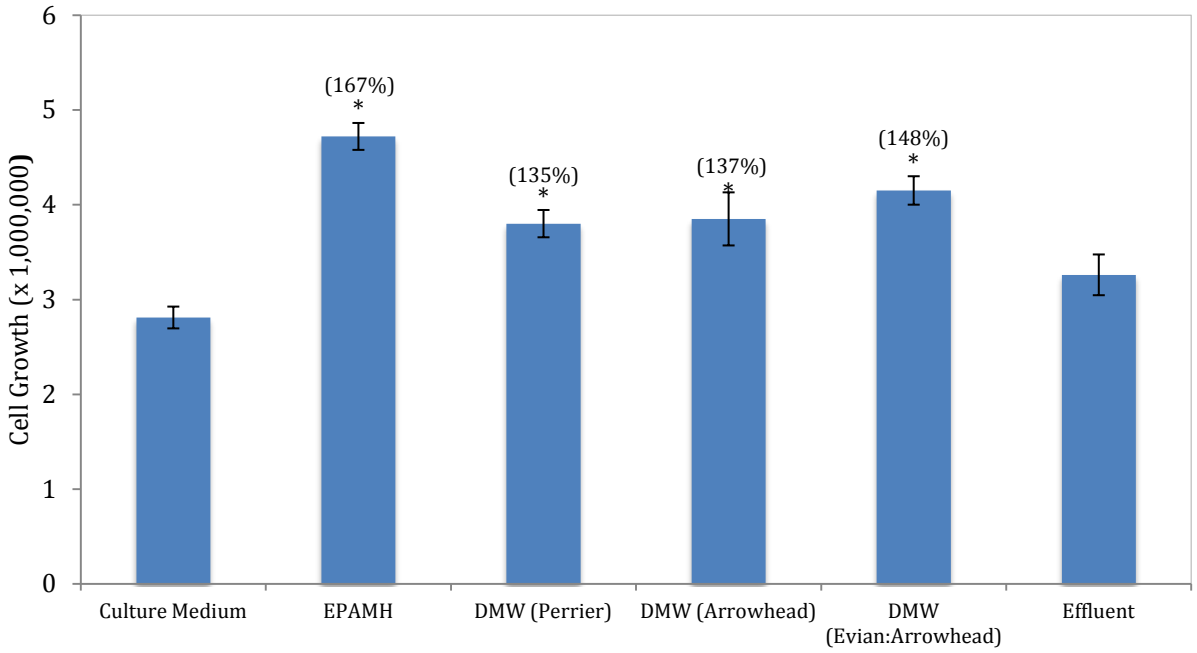

Figure 30. Comparison of Various *Selenastrum capricornutum* Laboratory Water Treatments to an Effluent Sample.

Table 13. Macro- and Micro-Nutrient Concentrations for Five Control Treatments Acceptable for *Selenastrum capricornutum* Testing

Macro- and Micro-Nutrient Concentration (µg/L)	Treatment				
	Culture Medium	EPAMH	DMW (Perrier)	DMW (Arrowhead)	DMW (Perrier:Arrowhead)
Bicarbonate	9,000	75,000	102,000	43,000	89,000
Boron	54	61	52	39	46
Calcium	1,300	17,000	38,000	6,600	22,000
Carbonate	<1,200	<1,200	<1,200	<1,200	<1,200
Chloride	6,000	11,000	12,000	6,200	7,900
Cobalt	0.7	<0.5	<0.5	0.6	<0.5
Copper	<1	<1	<1	<1	<1
Fluoride	<10	<10	<10	<10	<10
Iron	50	20	40	40	40
Magnesium	2,700	15,000	3,900	4,300	9,100
Manganese	130	120	110	120	140
Molybdenum	5	2	<1	<1	<1
Nitrate (as NO_3)	17,000	15,000	18,000	15,000	16,000
Phosphorous (total)	220	200	230	250	250
Potassium	600	3,300	700	1,800	1,200
Selenium (IV)	840	873	832	765	911
Sodium	9,800	37,000	12,000	12,000	12,000
Sulfate (as SO_4)	5,200	84,000	15,000	5,600	8,300
Zinc	<6	<6	<6	<6	<6

The conclusion of this study was that the various laboratory control waters that are used by laboratories can affect the outcome of the determination of the presence/absence of toxicity for POTW effluent samples. Clark et al (2011) recommended that laboratories use the algal culture medium as their laboratory control water since other acceptable control water options can result in stimulatory growth (i.e., 'false positives') due to the additional macro- and micro nutrients in the alternative controls.

5.1.2 Case Study 2 – Influence of Laboratory Water Medium on *Ceriodaphnia dubia* Inter-laboratory Variability

Sporadic toxicity to *C. dubia* reproduction was observed for discharge from a diamond mine in 2014 and 2015 (Pacholski et al., 2017). As there was a lack of correlation of the toxicity to water chemistry and there were no adverse effects in the zooplankton population in the receiving water, the facility was concerned that variability in the *C. dubia* tests could have resulted in a false positive. Variability in the testing was observed within individual samples, including IC_{25} values ranging from 56 to >100 percent (May 2014) and 42 to >100 percent (September 2014), but in all cases the test acceptability criteria were met. To further assess if *C. dubia* test variability resulted in 'false positives', a 'split-laboratory' comparison was conducted involving three accredited Canadian laboratories with split-samples from multiple sampling stations evaluated. The Canadian *C. dubia* chronic test method is very similar to the USEPA method, but does allow for bracketing the age of the less than 24 hour test organisms in 12-hour increments rather than the 8-hour increments required in the USEPA method. As discussed below, as long as organisms are less than 24 hours old, the 12-hour age bracketing requirement for the Canadian method is not expected to be the factor causing divergent test results in this study.

The laboratories all performed the testing in a similar manner (e.g., same dilution series, test conditions), but it was determined that the laboratories used different (but all acceptable within the method) control/dilution waters:

- Laboratory A used a moderately-hard reconstituted water supplemented (60 mg/L $MgSO_4$, 4 mg/L KCl, 96 mg/L $NaHCO_3$, 60 mg/L $CaSO_4$) with 2 µg/L vitamin B12 and 5 µg/L Na_2SeO_3; control hardness was 80 mg/L $CaCO_3$;
- Laboratory B used DMW (deionized water + 20% Perrier water, supplemented with 2 µg/L vitamin B12 and 9.5 µg/L selenium [from $Na_2SeO_4 \bullet 10\ H_2O$ stock solution]; control hardness was 82 mg/L $CaCO_3$; and
- Laboratory C used reconstituted/de-chlorinated municipal drinking water and distilled water (60 mg/L $MgSO_4$, 4 mg/L KCl, 96 mg/L $NaHCO_3$, 40 mg/L $CaSO_4$, 8 µg/L vitamin B12, 8 µg/L selenium [from $Na_2SeO_4 \bullet 10$ H2O stock solution]; control hardness was 130 mg/L $CaCO_3$.

Other parameters that differed among the laboratories were the size/material of test chambers (i.e., glass and plastic), type of reference toxicant material (NaCl and zinc), and the dates of sample receipt and test initiation (i.e., all tests were initiated within the maximum three-day holding time limit required by the Canadian testing method, which differs from the 36-hour hold time requirement for USEPA methodology); although the holding time is different from the USEPA protocol, this is not expected to result in appreciable differences between laboratory variability observed following the USEPA method since the laboratories tested the samples under similar minimal test conditions.

The results of the study are presented in **Figure 31**. All tests met the test acceptability criteria of 80 percent survival and 15 or greater offspring/female produced by surviving controls in their first three broods. Laboratory control reproduction ranged from 16.1 to 33.1 offspring in the first three broods. The control reproduction was lower for Laboratories A and C since there were one or more control replicates that produced few or no offspring; the authors noted that this is a less

desirable result since it increases replicate (i.e., intra-test) variability. Laboratory B exhibited better overall reproduction performance using DMW as their control/dilution water compared to the moderately-hard reconstituted waters used by Laboratories A and C.

The authors concluded, based on higher laboratory control reproductive performance, that DMW should be used for culturing *C. dubia* and as a control/dilution water, and further state that *"the use of the same dilution/control water by all laboratories in this comparison would have eliminated one of many potential sources of inter-laboratory variability."* However, it is important to note that the test method used in this study provides some laboratory flexibility for other test parameters (e.g., food quantity/type) that may also affect the outcome of 'split-laboratory' testing. As it remains to be demonstrated that more stringent requirements for laboratory control water alone will reduce variability in the chronic *C. dubia* test, at a minimum, this factor as well as others should be evaluated in future inter-laboratory comparability studies.

5.1.3 Factors Affecting Test Species

Test Design

Per the **Section 5.2** of this report, studies identified the *P. promelas* as the most consistent test among laboratories, followed by the *S. capricornutum* test, and finally the *C. dubia* test. It is likely that some of the reason for the greater consistency among laboratories for the *P. promelas* and *S. capricornutum* testing is based on the test design, as both species have multiple organisms per test replicate while the *C. dubia* test has a single organism per replicate (USEPA 2002). Tests with multiple organisms per replicate typically have greater intra-test and intra-laboratory precision.

Unique to the *C. dubia* test is the test termination criteria – tests are to be terminated when the 60 percent of the replicates in the control treatment have a third brood (maximum of 8 days). When even a single control replicate doesn't achieve three broods, the intra-test variability will increase; the intra-test variability can be increased further since up to four replicates in the control are not required to have a third brood prior to test termination. In addition, the intra-test variability can increase should any of the replicates in the effluent treatment(s) not have three broods. Should the test for one laboratory participating in a 'split-laboratory' evaluation have replicates that do not all have three broods, while the other laboratory has all replicates achieving three broods, it would be expected that the former laboratory would have a reduced capability to detect significant differences among treatments (i.e., they would have lower precision), while the latter would have a greater capacity to detect significant differences (i.e., they would have greater precision) following the standard hypothesis testing approach in USEPA (2002).

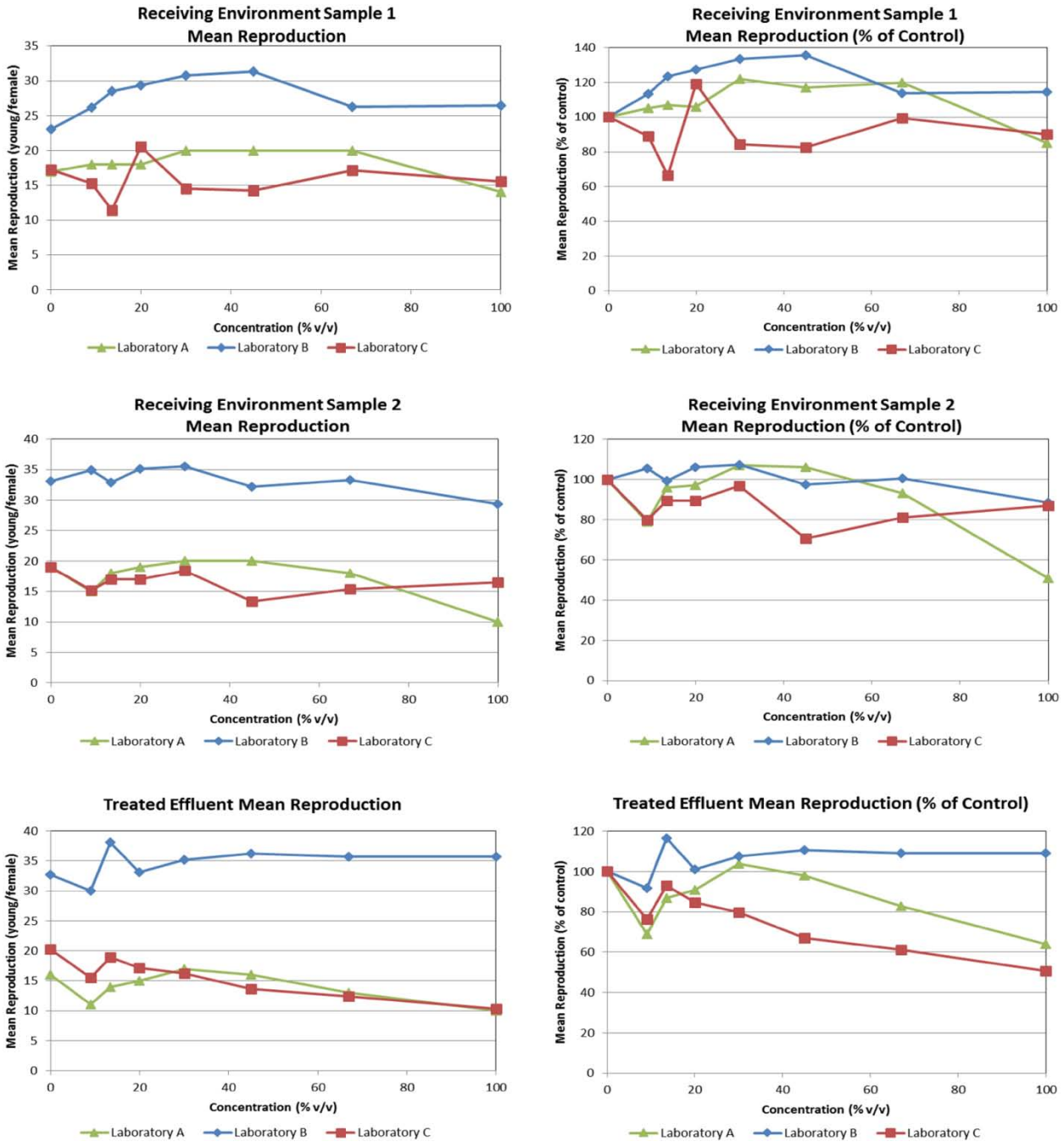

Figure 31. *Ceriodaphnia dubia* Inter-Laboratory Comparison Test Results for Receiving Water and Effluent Samples (Pacholski et al., 2017)

Sex Ratio

Sex ratios are not an issue for *S. capricornutum* since there is no sex for algae. Sex ratios are unlikely to be an issue for the *P. promelas* test, as it would be assumed that a 50/50 sex ratio would occur in cultures managed by organism vendors (i.e., typically the source of organisms used for testing by laboratories).

Sex ratios may be an issue in *C. dubia* testing, as the goal for laboratory cultures of this species is to maintain the organisms in an asexual (i.e., parthenogenesis) reproductive mode, where the females essentially clone themselves. When males are observed in *C. dubia* cultures (or tests), it is a sign that the culture is stressed, which can be caused by inadequate food (USEPA 2002, Section 13.6.16.9.3) and is generally associated with conditions of environmental stress (Pennack 1989). The occurrence of males in a healthy, well-maintained culture is rare (USEPA 2002, Section 13.10.9.3.1) and impacts the intra-test variability. The method appropriately requires the elimination of replicates with males from the reproduction counts, but this decreases the total number of replicates which correspondingly decrease the intra-test precision; this would be a factor that could also decrease inter-laboratory precision should results of two laboratories be compared where one laboratory had males in their testing.

Organism Age

The age of organisms might be considered a source of variability in testing. However, it is unlikely to be the case for the *S. capricornutum* test since the required use of four- to seven-day old cultures should result in inoculating the testing with cultures that are in log-phase growth.

Similarly, the *P. promelas* test requires the use of organisms less than 24-hours old if cultured in-house or less than 48-hours old if purchased from a vendor. Due to the effort required to culture *P. promelas*, most laboratories purchase these test organisms from vendors, so it is likely that most laboratories are initiating tests with test organisms less than 48-hours old.

The chronic *C. dubia* test requires the use of organisms that are less than 24-hours old, but all within an 8-hour timeframe. Cooney et al., (1992) evaluated a variety of parameters that could affect the chronic *C. dubia* test. They found organism age at test initiation was not critical as long as the organisms were less than 24-hours old and randomly distributed among replicates. Therefore, it is unlikely that the required age of any of the chronic test species is a significant source of variability.

5.2 INTER-LABORATORY VARIABILITY

All scientific methods have variability within repeated measures within a laboratory (i.e., via duplicate testing within a laboratory) and among laboratories (i.e., inter-laboratory variability). To support this study, the grey and published literature were searched for documents and studies that evaluated the inter-laboratory variability of the *S. capricornutum* growth, *C. dubia* reproduction, and *P. promelas* growth endpoint as performed under USEPA Methods 1003, 1002, and 1000, respectively (commonly known as the freshwater three-species tests). There have been a number of USEPA documents which evaluated test variability/precision, including:

- Technical Support Document for Water Quality-based Toxics Control (USEPA 1991);

- Understanding and Accounting for Method Variability in Whole Effluent Toxicity Applications Under the National Pollutant Discharge Elimination System Program (USEPA 2000); and
- Final Report: Interlaboratory Variability Study of EPA Short-term Chronic and Acute Whole Effluent Toxicity Test Methods, Vol. 1 & Vol. 2 (USEPA 2001).

The above references cited seven different inter-laboratory studies that were evaluated for this study, and are referenced in the figures below.

To provide as much supporting information as available in this study, additional sources of inter-laboratory variability were solicited from organizations supporting POTWs with NPDES permits. The Southern California Coastal Water Research Program (SCCWRP) performed an inter-laboratory variability study in support of the Southern California Stormwater Monitoring Coalition; although multiple test species were included in their study, only the chronic *C. dubia* test is applicable to this study. The Bay Area Clean Water Association Lab Committee (BACWA-LC), CVCWA, and LACSD were contacted to determine if they had access to inter-laboratory studies for the freshwater three-species tests. No submittals were provided by the BACWA-LC and the only inter-laboratory data obtained from CVCWA and LACSD were primarily limited to evaluations in which the results were compared between two laboratories. Since the comparison of results between two laboratories does not provide comprehensive information regarding inter-laboratory method variability, such data were evaluated in **Section 4.5**.

A major source of large-scale inter-laboratory data that were applicable to this study is the annual proficiency testing performed by laboratories under the Discharge Monitoring Report-Quality Assurance (DMR-QA) program. For chronic three-species testing, only the *C. dubia* and *P. promelas* tests are required for DMR-QA. Although laboratories are required to report select point estimate (i.e., IC_{25}) and NOEC, the State Water Board could not release the inter-laboratory data for the evaluation performed for this study because the program is authorized by the USEPA. Requests were submitted to USEPA Region IX for the IC_{25} data, which Region IX in-turn forwarded to USEPA Headquarters where the request currently stands and is still under consideration.

While the scientific literature includes a number of publications regarding method variability, most of these studies were addressed in broader evaluations of inter-laboratory variability (e.g., USEPA 1991, USEPA 2000, USEPA 2001). The measure of test method variability that is consistently in large inter-laboratory comparison studies is the coefficient of variation (CV, calculated as the standard deviation divided by the mean) typically using the EC_{25} or IC_{25}. Furthermore, most studies published in the early literature were performed prior to the 2002 update of the USEPA test methods.

5.2.1 *Selenastrum capricornutum* Inter-laboratory Variability

No *S. capricornutum* inter-laboratory variability data were presented in the early USEPA inter-laboratory studies (USEPA 1991, USEPA 2000). USEPA later evaluated inter-laboratory variability for the *S. capricornutum* growth endpoint (with EDTA in the nutrient amendment) among 11 laboratories that were provided with a reference toxicant, spiked effluent, and spiked

receiving water samples. As shown in **Figure 32**, the percent coefficient of variation (%CV) for IC_{25} ranged from 9 to 39 percent depending on the sample type.

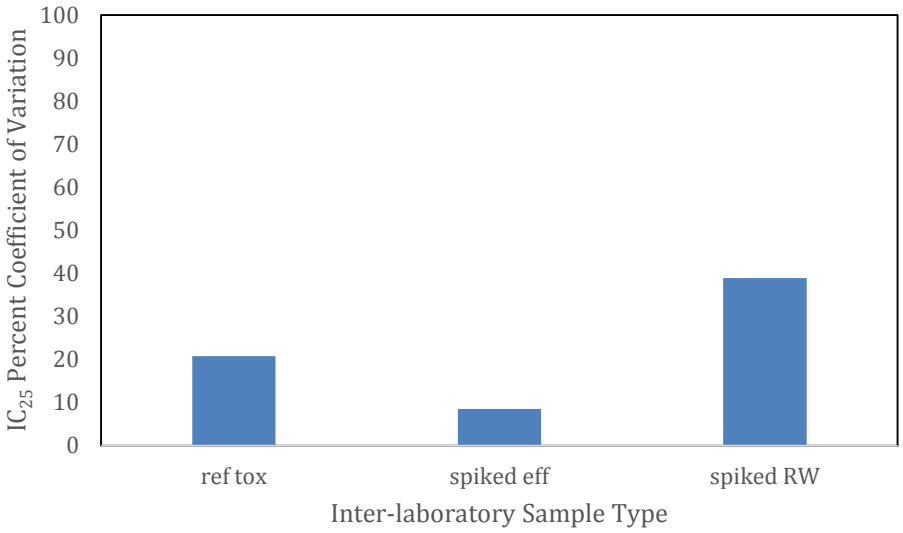

Figure 32. Inter-laboratory IC_{25} Percent Coefficient of Variation for the *Selenastrum capricornutum* Growth (USEPA 2001)

5.2.2 *Ceriodaphnia dubia* Inter-laboratory Variability

A number of inter-laboratory variability studies that included *C. dubia* occurred in the 1980s and 1990s with the number of participating laboratories ranging from 6 to 27. As shown in **Figure 33**, the IC_{25} %CVs ranged from 21 to 39 percent with no trend of improved test precision over this window of time. However, the information gained from the inter-laboratory studies performed in the late 1990s (i.e., Study 6 & 7, Figure 2) were used to adjust quality assurance requirements in the USEPA method, resulting in a new method manual (USEPA 2002).

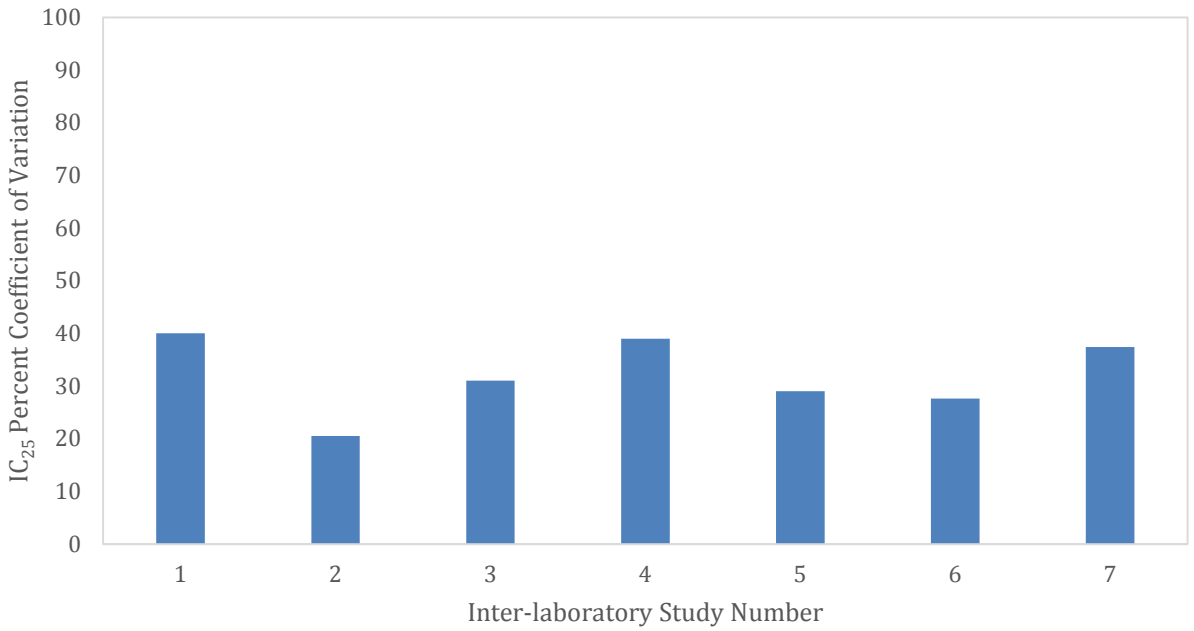

Figure 33. Inter-laboratory IC_{25} Percent Coefficient of Variation for the *Ceriodaphnia dubia* Reproduction (USEPA 1991, 2000, & 2001)

It would be expected that method variability would decrease as laboratories become more familiar with performing chronic toxicity tests, and following the improved QA requirements in the revised method manual. The only large inter-laboratory study to include the chronic *C. dubia* test since the adoption of the 2002 method manuals was performed by SCCWRP, and included the comparison of blanks, a copper-spiked lab dilution water, an artificial runoff sample, and a duplicates (Schiff and Greenstein, 2016). Nine laboratories successfully completed the first round of testing, which includes academic, municipal, and private sector laboratories that have been accredited for decades, and two nationally-accredited (i.e., National Environmental Laboratory Accreditation Program [NELAP]) laboratories. The IC_{25} %CVs ranged from 29 to 62 percent (**Figure 34**) during the first round of testing. The most consistent responses among the participating laboratories for *C. dubia* reproduction was for the artificial runoff sample, and the most variable results were obtained for the copper-spiked lab dilution water sample and the laboratory dilution water.

Due to the high inter-laboratory variability observed for this method, a second round of inter-laboratory testing was performed with six participating laboratories that agreed to ensure that certain method requirements (e.g., randomization, food type/feeding frequency) were standardized. However, there was no appreciable improvement in inter-laboratory precision (i.e., as measured by IC_{25} %CVs) and there was no marked improvement in the consistency of reproduction responses for the different sample types. The SCCWRP study identified that there was no clear relationship between feeding and water renewals and test variability, or hardness and test variability; the source of the variability was undetermined.

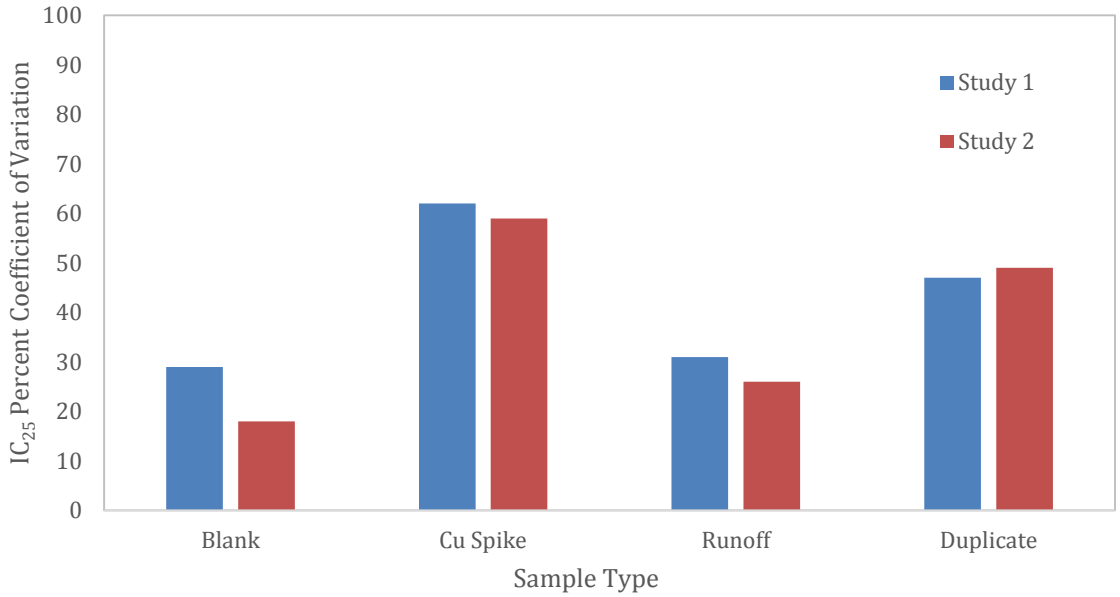

Figure 34. Inter-laboratory IC_{25} Percent Coefficient of Variation for the *Ceriodaphnia dubia* Reproduction (Schiff and Greenstein, 2016)

Although the SCCWRP study clearly leads one to conclude that additional research is needed to determine the source(s) of variability that caused such high inter-laboratory variability for the *C. dubia* test, the authors note that all laboratories met the test acceptability requirements, including internal positive and negative controls, and that most laboratories tended to produce internally consistent results when provided blind duplicate samples. In short, the laboratories internal quality control was good and their intra-laboratory variability (as assessed via duplicate samples) was good, but the inter-laboratory variability was high.

The variability observed in the SCCWRP study for the chronic *C. dubia* test is similar to variability observed in other studies of wastewater effluents, reference toxicants, and ambient water samples. Moore et al. (2000) evaluated laboratory variability for dilution water (i.e., blanks) among 16 laboratories, and observed a mean response of 16 percent effect and a standard deviation of 28 percent effect, which is comparable to the variability observed in the SCCWRP study for laboratory control water, which ranged from 16 to 27 percent effect, with a standard deviation of 19 to 27 percent effect. Observations of such low-level toxicity in blank samples from multiple studies are disconcerting and warrant further research.

5.2.3 *Pimephales promelas* Inter-laboratory Variability

A number of inter-laboratory variability studies that included *P. promelas* occurred in the 1980s and 1990s, with the number of participating laboratories ranging from 6 to 27; the IC_{25} %CVs ranged from 13 to 27 percent (**Figure 35**).

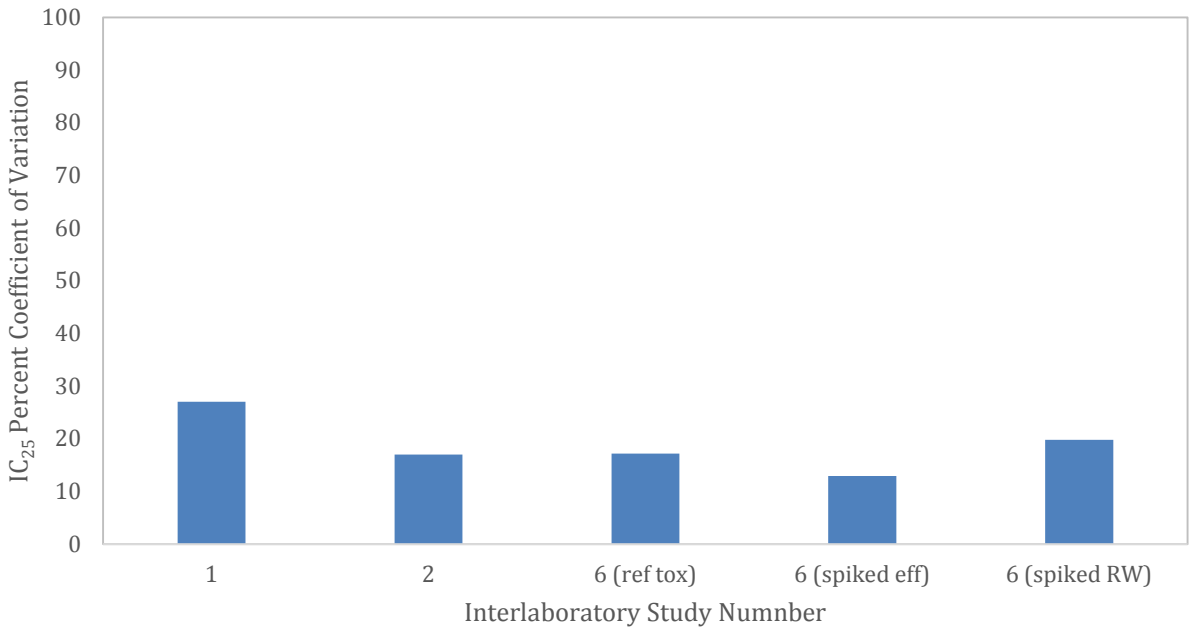

Figure 35. Inter-laboratory IC_{25} Percent Coefficient of Variation for the *Pimephales promelas* Growth (USEPA 1991 & 2001).

5.3 SUMMARY OF VARIABILITY IN CHRONIC TOXICITY TESTING

5.3.1 Intra-Chronic Toxicity Testing and Intra-Laboratory Variability

Although all laboratories in California are required to be accredited by the State Water Board's Environmental Laboratory Accreditation Program, the chronic toxicity test method provides flexibility on various factors ranging from test organism culturing (e.g., food, genetic variability) to microbial interferences to laboratory analyst training. These factors, among others, can impact the conclusions developed from chronic toxicity testing and the resultant findings on determining compliance with water quality standards.

5.3.2 Inter-laboratory Variability

Of the three chronic test species required to be tested by Central Valley POTWs for which inter-laboratory studies have been performed, the chronic *P. pimephales* growth exhibited the highest precision, followed by the *S. capricornutum* growth endpoint; the *C. dubia* reproduction endpoint exhibited the lowest precision even though there have been more inter-laboratory studies performed for this species and endpoint over a 30-year period of time. There were discussions during a toxicity testing method session at the Society of Environmental Toxicology and Chemistry (SETAC) – North America meeting in 2017 that a SETAC Issues Group (e.g., *Ceriodaphnia* Issues Group) should be formed to address the poor inter-laboratory precision observed for this test method, and that it should support potential future work by SCCWRP to better identify the sources and seek solution for the poor precision of this method. Similar SETAC groups were instrumental in identifying drivers for poor inter-laboratory precision for the 42-day *Hyalella Azteca* and life-cycle *Chironomus dilutus* sediment tests.

Section 6. Relationship Between Toxicity Testing and Aquatic Ecosystem Impacts

While WET testing is often used to assess final effluent from POTWs, it only provides a snapshot of the effluent and characterizes it as toxic or non-toxic isolated from its actual impact on the receiving waters and aquatic life. One of the stated goals of this study is to assess whether information exists, and to compile that which does exist, to identify levels of sub-lethal endpoint toxicity that do and do not correlate to measurable effects to aquatic life in receiving waters using bioassessment of receiving water data. As stated previously, the Basin Plan includes a narrative objective for toxicity. The Basin Plan continues:

> "*Compliance with this objective will be determined by analyses of indicator organisms, species diversity, population density, growth anomalies, and biotoxicity tests of appropriate duration or other methods as specified by the Regional Water Board.*"

If POTWs are unable to resolve the cause of chronic toxicity in effluent, other in-stream evaluations (e.g., indicator organisms, species diversity, population density) may need to be considered to assess if the effluent is causing toxicity in the receiving water and assess compliance with the narrative toxicity objective.

A literature review was conducted to summarize and discuss the general history of studies that have examined the link between toxicity tests and aquatic ecosystem effects, with a specific emphasis on levels of sub-lethal endpoint toxicity and characteristics that do and do not correlate to measurable effects to aquatic life, so that overall conclusions that summarize the state of the science can be made. The purpose is not to exhaustively detail and discuss every study that has examined the link between toxicity tests and aquatic life in the receiving water; therefore, not all potentially relevant studies are discussed, and the emphasis is on larger studies, review papers, and those that specifically relate to the goal stated above.

It is particularly challenging to discern from literature an answer to the stated goal (i.e., levels of sub-lethal endpoint toxicity that do or do not correlate with receiving water effects) for many reasons. Some of these reasons are related to study design, including, but not limited to, the following examples:

- Study designs vary widely;
- Not all characteristics of toxicity or bioassessment are reported in all studies;
- Statistical methods may lump lethal and sub-lethal endpoints which confounds interpretation of results; and
- Each physical site is unique and is usually affected by multiple factors beyond effluent quality.

In addition, because there is substantial complexity and natural variability in aquatic ecosystems, toxicity detected in the lab may degrade rapidly in the environment, and may do so at different rates at different times or locations. Moreover, dilution may not be adequately accounted for, bioassessment and toxicity testing are often asynchronous, exposure times can vary between the lab and the field, bioassessment methods vary, and there may be bias in the selection of sites

included in the study. Another set of reasons why the stated goal is particularly challenging is related to chronic toxicity testing laboratory variability and data quality including the following examples:

- Some studies do not always meet test acceptability criteria;
- The laboratories did not always perform reference toxicant tests or reference toxicant test results demonstrate at various times that the test organisms may be outside their typical sensitivity for such tests;
- Tests run within and between studies have varying levels of statistical power (e.g., PMSD); and
- 'Split-laboratory' testing may not have resulted in agreement between laboratories.

Finally, concerns related to bioassessment protocol data quality are also present. For example, sampling technique may vary between members of a sampling team or between different teams, samples may not be representative of the entire ecological community, and laboratory subsampling may not be representative of the entire sample.

Because of the reasons cited above, it is generally only possible, given the available data, to make qualitative statements regarding the link between a chronic toxicity test result and the likelihood of that result predicting ecological effects than it is to make a determination of a quantitative level of toxicity. For example, there seems to be consensus that "…biological responses, as all measurements, are less reliable near detection limits." False positives: "…are of greater concern in situations where surface water [or] effluent toxicity is relatively low and near detection limits. The ability to reliably detect biological community impairments when the concentrations of toxic chemicals are near the effect thresholds is difficult; detection of such impairments also will be obscured by the complexity and natural variability in aquatic ecosystems" (de Vlaming and Norberg-King, 1999). However, what "near the effect thresholds" or "near detection limits" means is up for debate, and will vary depending on the site, study design, data quality, and other factors.

In the 1980s, USEPA sponsored eight separate studies, deemed the Complex Effluent Toxicity Testing Program (CETTP), which included studies of 80 sites in 8 different watersheds (USEPA 1991). Data from these studies, as a group, were subsequently analyzed by Dickson et al. (1992) and Marcus and McDonald (1992). USEPA's Technical Support Document for Water Quality Based Toxics Control (TSD; USEPA 1991) cited the CETTP studies, which together reported a qualitative correlation[9] between *P. promelas* growth and *C. dubia* reproduction, and downstream effects on fish, invertebrates, and periphyton as support for toxicity-based water quality control. The CETTP studies were criticized by Parkhurst et al. (1990) and Marcus and McDonald (1992) for, among other shortcomings, selecting sites with high instream toxicity and known biological impacts. The TSD responded to some criticisms made by Parkhurst et al. (1990), and the criticisms were again addressed in de Vlaming and Norberg-King (1999). For example, in response to the specific criticism of site-selection bias, de Vlaming and Norberg-King (1999)

[9] Presence of toxicity correlated with presence of downstream effects, though the correlation was not quantitative.

stated that "This criticism has merit and should be considered when evaluating the CETTP data. Design of the CETTP studies was not perfect from a statistical analysis standpoint…limited funds and other resources require regulatory agencies to focus on areas where there are likely to be environmental problems so there can be remediation and restoration" (1999). Despite the responses, the fact remains that due to the selection of sites with high instream toxicity and known biological impacts, it is uncertain whether results from these studies can be generalized to all POTWs and receiving waters.

de Vlaming and Norberg-King (1999) aspired to provide "…a comprehensive review to evaluate the reliability of single species (also referred to as indicator species) toxicity test results in predicting aquatic ecosystem impacts, also known as the ecological relevance of laboratory single species toxicity tests." The authors examined differences between ambient waters and effluents, evaluated published criticisms of single species tests, and evaluated whether laboratory toxicity tests are more sensitive than natural populations. It should be noted that the focus of their review was not specifically on chronic toxicity tests, but on single-species toxicity tests in general, which also includes testing of ambient water. de Vlaming and Norberg-King (1999) concluded that: "A preponderance of evidence reveals that laboratory single species toxicity test results are reliable qualitative predictors of aquatic ecosystem community impacts." This conclusion has not been universally accepted or confirmed with more recent study, as will be discussed later in this section. The following paragraphs from de Vlaming and Norberg-King (1999) in particular have great significance with regards to this study:

> "The reliability with which single species toxicity test results predict biological community responses relates to several factors. One major factor was addressed by Dickson et al. (1992); they observed that when effluent or ambient water toxicity is relatively low or when impacts on aquatic ecosystems are moderate it will be difficult to establish a relationship between toxicity and instream ecological responses. The strength of the predictive capacity of -single species test results is substantially enhanced when the test is performed with ambient water (e.g., as compared to effluent) and with higher magnitude toxicity in the sample. Chapman et al. (1987) came to a similar conclusion regarding magnitude of toxicity in relation to sediment tests. We appear to be approaching consensus that when significant lethality (and in the case of effluents, assuming accurate dilution has been considered) is seen in toxicity tests there is a very high potential of aquatic ecosystem impairment. As this connection is accepted, we continue to struggle with the idea that sub-lethal effects on indicator species can result in detectable adverse ecosystem responses."
>
> "A convincing relationship has been established between ambient water toxicity (as manifested by single species tests) and biological community responses, but has such connection been authenticated between effluent toxicity and instream impairments? The effluent-biological community link has not been as thoroughly investigated. Nonetheless, in several recent studies… as well as the CETTP and associated studies, where effluent toxicity was assessed, a reliable qualitative estimate of instream biological effects was obtained. This relationship was most evident when flow and dilution of the receiving water were effectively estimated and when environmental exposure duration was matched (or account for) by laboratory toxicity test duration."

> "The predictive power of single species tests is substantially enhanced when ambient water, as compared to discharge, is tested and when higher magnitude-toxicity exists; reliability is also improved when exposure patterns in natural ecosystems are matched or accounted for and, in the case of effluents, when realistic estimates of dilution are taken into account."

Thus, de Vlaming and Norberg-King (1999) defend the utility and effectiveness of single species toxicity tests as a general tool, but conclude that: 1) ambient water toxicity testing is a better predictor of ecological effects than effluent toxicity testing, 2) that higher magnitudes of toxicity/lethal endpoints are more predictive of ecological effects than are sub-lethal endpoints, and 3) that effluent testing becomes a better predictor of ecological effects for when dilution is accounted.

In contrast to de Vlaming and Norberg-King (1999), Chapman (2000) assessed the status of toxicity tests specifically in 1999-2000 relative to their general purpose of "…identifying, characterizing, and eliminating toxic effects of effluents on aquatic resources." The review discusses many factors, including variability of the tests, differences in species between the laboratory and the field, and differences between the laboratory and receiving water environment. Instances of overprotection, underprotection, and uncertain level of protection are discussed. Chapman (2000) reviews and discusses many of the same studies discussed in this literature review, and comes to the conclusion that: "Comparisons to field conditions indicate that WET tests are not reliable predictors of effects or lack of effects in the receiving environment." He then goes on to discuss the concept of "independent applicability", in which toxicity test results can be considered in isolation from information on the receiving environment, and makes the argument that: "Whole effluent toxicity tests are only the first stage in a risk assessment and as such identify hazard, not risk. Identification of risk requires discarding the concept of independent applicability," and "Whole effluent toxicity tests are appropriate for identifying and delineating effluents of concern (i.e., hazard ranking)…Ideally, whole effluent toxicity tests would serve as flags to indicate the need for a more formal risk-based approach . . . [they] can be thought of as a low budget alternative to the risk assessment framework as applied to an individual point source. . . . The extent to which whole effluent toxicity serves as an effective indicator of ecological risk remains open to debate.'' Chapman (2000) argues that independent applicability should not be used in regards to WET tests, and states that "because WET tests only serve to identify hazard, not risk, they should not be used alone..."

In order to assess a more comprehensive data set, and to avoid selecting sites with known biological impairments (i.e., to avoid the problems the CETTP studies had), Diamond and Daley (2000) compiled and used a database of 250 dischargers across the US (including four dischargers in California) to examine the relationship between *C. dubia* and *P. promelas* WET results (as opposed to ambient toxicity test results) and instream biological condition. In general, WET test results were well associated with instream biological condition only when all of the following were true.

- The effluent comprised 80 percent or more of stream flow under design conditions.[10] More dilute effluent exhibited substantially less agreement between WET and instream conditions.
- Instream habitat quality was characterized as fair to good.
- At least three WET tests had been conducted.
- A test failure rate (i.e., toxicity was detected via the statistical test used) of at least 25 percent had occurred. Test failure rate was defined using thresholds based on LC_{50} values (i.e., LC_{50} was less than 100 percent effluent) for acute WET tests and/or NOECs for chronic WET tests (i.e., NOEC/IWC was less than 1).

Single WET test failures were unrelated to stream impairment, and effluents that made up less than 20 percent of the stream flow under design conditions had a low probability of being associated with instream biological condition. The study found that *P. promelas* endpoint toxicity was positively correlated with stream impairment, and *C. dubia* acute and chronic survival was inversely correlated with stream impairment (i.e., the opposite of being predictive of instream biological condition). *C. dubia* reproduction showed no relationship with instream biological condition. The authors also examined the correlation of inter-test variability (i.e., how variable were results of repeated tests over time at the same facility) to instream condition, and found that only when inter-test variability was low (CV of the chronic *P. promelas* lowest observed effect concentration [LOEC] was less than 20 percent) and toxicity was routinely present were results predictive of instream condition. If inter-test variability was high, no correlation between results and instream impairment was present. Finally, Diamond and Daley (2000) concluded "Effluent dilution was the strongest factor affecting relationships between WET and observed biological conditions."

On the whole, in the meta-data analysis conducted by Diamond and Daley (2000), approximately 47 percent of the time there was disagreement between WET results and bioassessment results (i.e., WET failed but stream was unimpaired (22.8 percent), or WET passed and stream was impaired (23.9 percent)). The authors state that: "The above results suggest two important ramifications for WET programs as currently practiced. First, 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 test failure rate (≥90%). Second, there is roughly a 20% probability that impairment may be observed instream even though WET did not indicate reasonable toxicity potential, depending on which type of tests were conducted."

In Diamond et al. (2008), the authors remarked that the study discussed above that examined the large database of dischargers (i.e., Diamond and Daley 2000) was somewhat inconclusive, based on missing data, outdated WET test methods, or insufficient quality of bioassessment methods. Diamond et al. (2008) sought to shed greater light on the issue by introducing several measurement quality objectives that specified desired precision, bias, and sensitivity of the methods. Six facilities that all had design effluent concentrations greater than 60 percent of the

[10] Effluents that made up 80 percent or more of the stream flow for reported 7Q10 conditions "had a significant relationship between WET and instream biological condition if WET compliance was defined as passing between 75 and 85 percent of all WET tests conducted" (Diamond and Daley 2000).

stream flow were used. Splits, duplicates, and blind positive and negative controls were employed and all facilities performed quarterly three-species testing. Macroinvertebrate, fish, and periphyton bioassessments were conducted upstream and downstream of each facility. The study showed that using the IC_{25} endpoint as opposed to the NOEC endpoint resulted in fewer false positives, less toxicity, and less "failed"/noncompliant tests. The algae test was most often found to be toxic, but also showed very high (the greatest) inter-laboratory variability and the most false positives. The authors stated that: "Overall, WET test results exhibited few relationships with bioassessment results even when accounting for actual effluent dilution. In general, neither frequency of WET noncompliance nor magnitude of toxicity in tests was significantly related to differences in biological condition upstream and downstream of a discharge." For example, the site showing the highest toxicity to *S. capricornutum* (14.4 TU_c based on actual dilution) showed no effects instream. Repeatability of tests at the same laboratory was good, but inter-laboratory variability was not. *P. promelas* and *C. dubia* survival endpoint results were similar across laboratories for all facilities, but sub-lethal endpoints showed greater than 35 percent difference in NOEC or IC_{25} for 2-5 of the 6 facilities (depending on species). The authors stated that "it appears that compliance with test acceptance criteria, while necessary and important, is not sufficient for evaluating laboratory performance and data quality for freshwater chronic WET tests." Regarding this statement, the authors argue that data and measurement quality objectives that include use of blind positive controls, negative controls, and split sampling would improve predictions beyond just evaluating test acceptability criteria. However, the authors also caution that: "This case study demonstrated ways in which WET results could be compared with biological condition, but the results observed here should not be considered a definitive assessment of relationships between the 2 types of monitoring. Because this was a pilot study, it was not designed to determine definitive relationships between WET results and biological condition."

6.1 CASE STUDY: CITY OF WOODLAND *SELENASTRUM CAPRICORNUTUM* TOXICITY REDUCTION EVALUATION

The City of Woodland (City) was engaged in a TRE for *S. capricornutum* at its Water Pollution Control Facility (WPCF) from 2009 to 2015. Toxicity to *S. capricornutum* was intermittent and low-level. The City investigated various potential contributors to toxicity, and undertook actions including, but not limited to, the following: extensive Facility Performance Review, five Phase I TIEs as well as several hold-time studies to determine stability of the toxicant(s), Toxicity Source Evaluations including testing of effluent at different points in the treatment process, bioassessment of Tule Canal (the receiving water), and assessment of dilution available in Tule Canal. The residual effects of irradiating effluent in the WPCF ultraviolet light disinfection system was identified as the primary cause of growth inhibition to *S. capricornutum* in WET tests. With the exception of a brief period in 2009, all bioassay results that were toxic had a TU_c of 2 (14 to 75 percent growth inhibition in the 100 percent effluent relative to laboratory control water), resulting in a NOEC of 50 percent effluent. Since receiving water samples from upstream of the WPCF discharge rarely indicated toxicity, this meant that at greater than 1:1 river:effluent dilution ratios, toxicity would not be expected. A dilution analysis was performed that showed that Tule Canal nearly always had enough flow such that dilution was sufficient to eliminate concerns of potential toxicity in the receiving water, based on bioassay results. Rapid bioassessment of the benthic macroinvertebrate (BMI) community in Tule Canal confirmed this, in that there was no significant difference in community structure and composition upstream

versus downstream of the outfall. In summary, chronic toxicity trigger exceedances had no detectable impairments of beneficial uses in the receiving water. Findings from the TRE provided sufficient justification for the Central Valley Water Board to adjust the NPDES permit chronic toxicity trigger from 1 to 2 TU_c for *S. capricornutum* to account for available dilution. Consistent compliance with the 2 TU_c trigger has been achieved, resolving the TRE.

6.2 CASE STUDY: EL DORADO IRRIGATION DISTRICT *CERIODAPHNIA DUBIA* TOXICITY REDUCTION EVALUATION

The El Dorado Irrigation District was engaged in a TRE for *C. dubia* at its Deer Creek Wastewater Treatment Plant (DCWWTP) from late 2015 to late 2016. Samples identified as toxic for this TRE typically ranged 1.3 to 2 TU_c (25 to 50 percent reproduction inhibition in the 100 percent effluent relative to laboratory control water). The TRE included the following major activities: facility performance review and evaluation, *C. dubia* bioassay testing including split sampling and testing at multiple laboratories, and BMI bioassessment of Deer Creek (DCWWTP's receiving water).

'Split-laboratory' testing performed among three separate laboratories provided an indication that the effects were intermittent, low-level, and within the variability of the test (i.e., not detected across multiple laboratories during the split tests).

A survey of Deer Creek's BMI community upstream and downstream of the DCWWTP discharge location was conducted. The BMI survey conducted in July 2016 for Deer Creek indicated a fully functioning BMI community upstream and downstream of the discharge, which was supported by the high taxonomic richness of the downstream sites (similar to upstream sites), similar presence of EPT taxa among reaches, Shannon Diversity index values in the expected range for a small foothill creek, a high degree of community similarity of downstream to upstream sites (i.e., Sørenson's QS index values), and all expected functional feeding groups present at all locations. Variability in the BMI community observed between sites located upstream and downstream of the discharge was explainable by changes in physical creek substrate and habitat, as evidence by the loss of the shredder community and bryophytes, which are specialists that live and feed exclusively on aquatic mosses. The range of bioassessment index values and taxa observed between upstream and downstream stations were similar to the range of values and taxa obtained for surveys of Deer Creek conducted in 2007, 2008, and 2009 when the facility was not observing greater than 1 TU_c WET test results.

The finding that the DCWWTP effluent discharge was not adversely impacting aquatic life in the creek, coupled with no indications from facility performance review, effluent quality, or plant operations supported the conclusion that the greater than 1 TU_c bioassay results intermittently observed during accelerated testing and the TRE were primarily a function of inter-laboratory variability in conducting the chronic *C. dubia* test. Based on these findings, the El Dorado Irrigation District concluded the TRE.

6.3 LITERATURE REVIEW SUMMARY

A number of common themes were identified from the body of studies discussed above. These themes are listed below.

1. In general, it does not appear that WET test results are reliable predictors of effects or lack of effects in the receiving water environment (Chapman 2000, Diamond 2000, Diamond et al. 2008). Specifically, intermittent and low-level toxicity, as measured by sub-lethal endpoints, does not appear to be a reliable predictor of receiving water impairments. Some studies, in particular the CETTP studies, have demonstrated a qualitative correlation between *P. promelas* and *C. dubia* ambient toxicity tests and instream biological condition, but there is considerable debate as to whether these studies are representative of effluents and their receiving waters in general.
2. In general, ambient water toxicity testing better represents biological condition of the water body than effluent testing, and higher magnitudes of ambient toxicity are better correlated with biological effects (de Vlaming and Norberg-King 1999, Dickson 1992).
3. WET testing is better representative of instream biological effects when dilution is considered and when higher magnitudes of toxicity are present after considering dilution (de Vlaming and Norberg-King 1999, Diamond et al. 2000, City of Woodland).
4. Higher frequencies and magnitudes of WET toxicity can be generally better correlated with biological effects in a water body (Dickson 1992, de Vlaming and Norberg-King 1999, Diamond et al. 2000).
5. There is no consensus on which WET test species provide the best predictions of biological condition in the receiving water. Different studies have reached different conclusions on this matter (Diamond and Daley 2000, Diamond et al. 2008).
6. Because biological responses measured in WET tests are considered less reliable near test detection limits (de Vlaming and Norberg-King 1999), predictions of biological effects in a waterbody based on WET testing will be improved when laboratory performance and data quality for freshwater chronic WET tests is evaluated with measurement quality objectives that include the use of split test evaluation, blind positive control testing, blind negative control testing, and reference toxicants (Diamond et al. 2008).

A goal of this study was to assess whether information exists, and to compile that which does exist, to identify levels of sub-lethal endpoint toxicity that do and do not correlate to measurable effects to aquatic life in receiving waters using bioassessment of receiving water data. To this end, a single, specific "level" of sub-lethal endpoint toxicity (i.e., magnitude of effect in WET tests) that correlates well to measurable effects in aquatic life could not be identified, largely because of point #1 above (in general, sub-lethal WET tests results are not reliable predictors of effects or lack of effects in the receiving environment). Furthermore, this was not specifically a study goal of any of the studies reviewed, and the existing studies are sometimes contradictory. Although it does not appear as if a single, well-defined threshold level can be identified, criteria that are associated with improved correlation between WET test results and measurable effects on aquatic life can be identified. For example, as stated above, Diamond and Daley (2000) found that WET results for *P. promelas* correlated with instream impairments when:

1. The effluent comprised 80 percent or more of stream flow under design conditions;
2. Instream habitat quality was characterized as fair to good;
3. At least three WET tests had been conducted; and
4. A test failure rate (i.e., toxicity had been detected) of at least 25 percent had occurred.

These are criteria that, when met, improved correlation between WET results and instream biological condition, but meeting these criteria did not guarantee predictive ability. To this end, a

set of criteria, based on the literature reviewed herein, is summarized and proposed below. It can be reasonably assumed, though not guaranteed, from this study that the more criteria that are satisfied below, the greater the hazard posed by the effluent to instream biological condition.

1. Test failure (i.e., toxicity) is defined in terms of the IC_{25}, rather than the NOEC.
2. A test failure rate (i.e., toxicity has been detected) of at least 25 percent has occurred.
3. Effluent comprises of at least 80 percent of the stream flow under design conditions, or effluent dilution within the receiving water has been accounted for.
4. Laboratory performance and data quality of WET tests has been evaluated using measurement quality objectives that include the use of 'split-laboratory' tests, blind positive controls, blind negative controls, and/or reference toxicants, and WET tests results do not appear related to inter-laboratory variability.
5. Downstream ambient toxicity testing has been performed and has indicated toxicity.
6. Inter-test variability over time is low (less than 20 percent CV between IC_{25} of multiple consecutive tests) and toxicity is routinely present.

This page is intentionally left blank.

Section 7. Draft Conceptual Model

A fundamental principal of toxicology is the association of increased effect, such as impairment of reproduction, with increased concentration, or dose, of toxicant. Typically, it assumed that a causal relationship exists between the concentration of a contaminant and a measured response in the organisms. For the chronic WET tests, the measured response could be in the growth of *S. capricornutum*, survival and reproduction of *C. dubia*, or the survival and growth of *P. promelas*. The classic concentration response would be considered a sigmoidal shaped curve, in which the response in the organism increases as the contaminant concentration increases, with more severe effects (e.g., acute survival) typically occurring at the higher concentrations and less severe responses (e.g., growth and reproduction) occurring at lower concentrations.

The challenge in designing a conceptual model specifically for low-level effects on a contaminant-based concentration response curve is that it would be expected that the same drivers that could cause a low-level effect could cause a higher-level effect based simply on the concentration of the contaminant. For this reason, the conceptual model designed for this study includes all potential drivers for toxicity (low-level or higher level), with many of the drivers in the conceptual model having been identified in TRE studies performed for Central Valley POTWs.

The toxicity conceptual model, described below, is divided into three elements: POTW drivers, testing laboratory drivers, and the environmental drivers.

7.1 POTW DRIVERS OF TOXICITY

Sources upstream of the POTW have been identified in Central Valley POTW TREs as potential causes of toxicity. These include industrial users (e.g., categorical, significant), collection system maintenance and operation activities (e.g., root control, slip lining), severe infiltration and inflow (I/I), and source water (e.g., ion imbalance, hardness). Other possible upstream sources include domestic and commercial dischargers, particularly as they may relate to disposal of unused or unwanted chemical products. At the POTW, operations, maintenance, and construction have been key sources of toxicity. In addition, sample collection issues (e.g., microbial growth in compositor tubing used to collect the effluent samples) have been identified as sources of observed toxicity. A number of dischargers are required to collect background receiving water, which is to be evaluated alongside the effluent for toxicity, and in some cases used as the control/dilution water for comparison to the effluent. Background receiving waters have been observed as causing:

- Stimulatory responses to *S. capricornutum* and *C. dubia*, presumably due to the presence of additional nutrients or food in the receiving water sample. If used as the control/dilution water, this can result in a ‘false positive’ for toxicity when the effluent is compared to the laboratory control water;
- Similarly, a ‘false negative’ result for toxicity can be observed based on a comparison of the effluent to a receiving water sample, where the effluent outperforms the receiving water, but simultaneously is identified as toxic when compared to a laboratory control

water, as in cases where the receiving water itself would be considered toxic when compared to the laboratory control;

- Pathogen-/microbial-related issues affecting *C. dubia* and *P. promelas* tests. Such observations are test method-specified interferences, and require further test design adjustments (e.g., additional replicates for the *P. promelas* test) or sample manipulations (e.g., microfiltration, antibiotics) to remove the test interference; and/or
- Toxicity in the *S. capricornutum, C. dubia*, and *P. promelas* tests indicating that the receiving water upstream of the effluent discharge was toxic, which would invalidate the evaluation of effluent toxicity if the receiving water was used as the control/diluent in the test.

7.2 TESTING LABORATORY DRIVERS OF TOXICITY

A number of laboratory issues can affect the outcome of the toxicity test. These sources have been categorized in the conceptual model into laboratory experience and expertise, organism quality and sensitivity, test interferences, and test design.

The experience and expertise of the laboratory can affect the toxicity test, as well as the quality and thoroughness of the quality assurance/quality control (QA/QC) program implemented in the laboratory; it is critical that laboratories have a good training program to ensure that their staff perform each method in a consistent manner, and that all staff adhere to the standard protocols. An example as to how this can affect the presence/absence of toxicity would be laboratory staff that do not perform accurate counts of the small neonate offspring in the *C. dubia* test.

Organism quality and sensitivity can affect the outcome of toxicity tests. Most laboratories do not culture their *P. promelas*, but rather purchase these organisms from vendors, and the organisms can exhibit shipping stress (e.g., handling of the shipping container, seasonal fluctuations in shipping temperatures). Culture health, whether cultured in-house or at a vendor, can affect the outcome of the toxicity testing; this can be assessed through performance of reference toxicant testing performed by the laboratory and by adherence to performance measures in the USEPA manual (2002) provides considerable flexibility in the type, quantity, and quality of food used in the *C. dubia* and *P. promelas* tests. Poor quality food can result in cultures that are of poor quality, which can in turn result in hypersensitivity to toxicant stress.

Multiple sources of test interferences can occur, including pathogen/microbial, test scoring, unusually high control response, and testing errors. Laboratory staff must be experienced in identifying, avoiding, and rectifying such interferences, otherwise test results may falsely identified an effluent as toxic rather than appropriately qualifying the test results as being due to test interference. Test scoring errors can occur due to improper counting of test organisms. An example as to how this can affect the presence/absence of toxicity would be if laboratory staff undercount the small neonate offspring in the *C. dubia* test. Another test scoring issue identified by laboratories has been ‘plating’ observed in some *S. capricornutum* tests, in which some of the algae appear to stick to the test chamber. As the USEPA (2002) indicates that all algae in the flasks are to be counted, laboratories should re-suspend and algae that have ‘plated’ in the test chambers to avoid undercounting the organisms. Finally, testing errors (e.g., incorrect organism age, water, food, etc.) can result in identifying samples as toxic when in fact they may not have been had the proper testing protocol been used.

Test design elements, including the number of replicates, statistical methods, and control/dilution medium can influence the determination of toxicity. Each test method requires a minimum number of test replicates (i.e., four for *S. capricornutum* and *P. promelas*, ten for *C. dubia*), and the NPDES permits specify the statistical requirements (e.g., NOEC in the Central Valley). Increased replication improves test precision, and can result in increased test sensitivity, while loss of replicates due to laboratory errors (e.g., spilling a test replicate) can decrease test precision, either of which can affect the outcome of the statistical analyses. As discussed in earlier sections of this report, USEPA (2002) provides considerable flexibility in which control/dilution water is used by the laboratory, and the NPDES permits can provide further guidance as to which control/dilution water is used for statistical comparisons to the effluent treatment(s). As noted previously in this report, the selection of control waters for the *S. capricornutum* test that produce stimulatory responses in the control (due to extra nutrients when compared to other control waters) can result in a finding of toxicity while another control medium may not identify the effluent as toxic. Similarly, the use of a moderately-hard water control treatment for comparison to a soft water effluent may indicate the effluent is toxic, while a soft water control may not.

7.3 ENVIRONMENTAL DRIVERS OF TOXICITY

Based on the literature review presented in **Section 6**, there are several issues that affect the level of hazard posed by the effluent to the receiving water. These sources have been categorized in the conceptual model as laboratory variability, instream conditions, and observed WET results. The level of hazard posed by the effluent is greatest when toxicity is observed and when criteria derived from the literature that are associated with improved correlation of WET results with instream biological condition are present.

Laboratory variability includes organism, intra-laboratory, and inter-laboratory variability. In general, the lower the variability observed in WET results, the higher likelihood that WET results will correlate with receiving water biological condition. The nature of observed WET results influences the level of hazard posed by the effluent. Items such as the frequency of observed toxicity, magnitude of WET results, and variability in the observed IC_{25} over time have been included. The literature review found that higher magnitudes (de Vlaming and Norberg-King 1999, Dickson 1992) and higher frequencies of WET toxicity are generally better correlated with biological effects in a waterbody (Diamond et al. 2000; Diamond et al. 2008).

Instream conditions include habitat condition, effluent dilution, and presence/absence of downstream ambient toxicity. For example, literature indicates that better agreement between WET results and instream biological condition is observed when instream habitat quality is characterized as fair to good (Diamond et al., 2000). Further, in general, ambient water toxicity testing better represents biological condition of water bodies than effluent testing (de Vlaming and Norberg-King 1999, Dickson 1992). Finally, WET testing is more representative of instream biological effects when dilution is considered (de Vlaming and Norberg-King 1999, Diamond et al. 2000).

7.4 DRAFT CONCEPTUAL MODEL

Section 4 of this report summarized an analysis that was conducted to characterize chronic toxicity test results that were observed in discharges from Central Valley POTWs. **Section 5** of this report summarized an evaluation of specific variables during chronic toxicity testing and within test species and organisms that can affect the outcome of a chronic toxicity test. **Section 6** of this report summarized the state of the science regarding the correlation between WET results and instream biological condition. In this section, the information developed in **Sections 4**, **5**, and **6** are blended to develop a draft conceptual model that provides a general guide on factors to consider when conducting chronic toxicity tests and follow-up TRE/TIE testing. The draft conceptual model is presented in **Figure 36**.

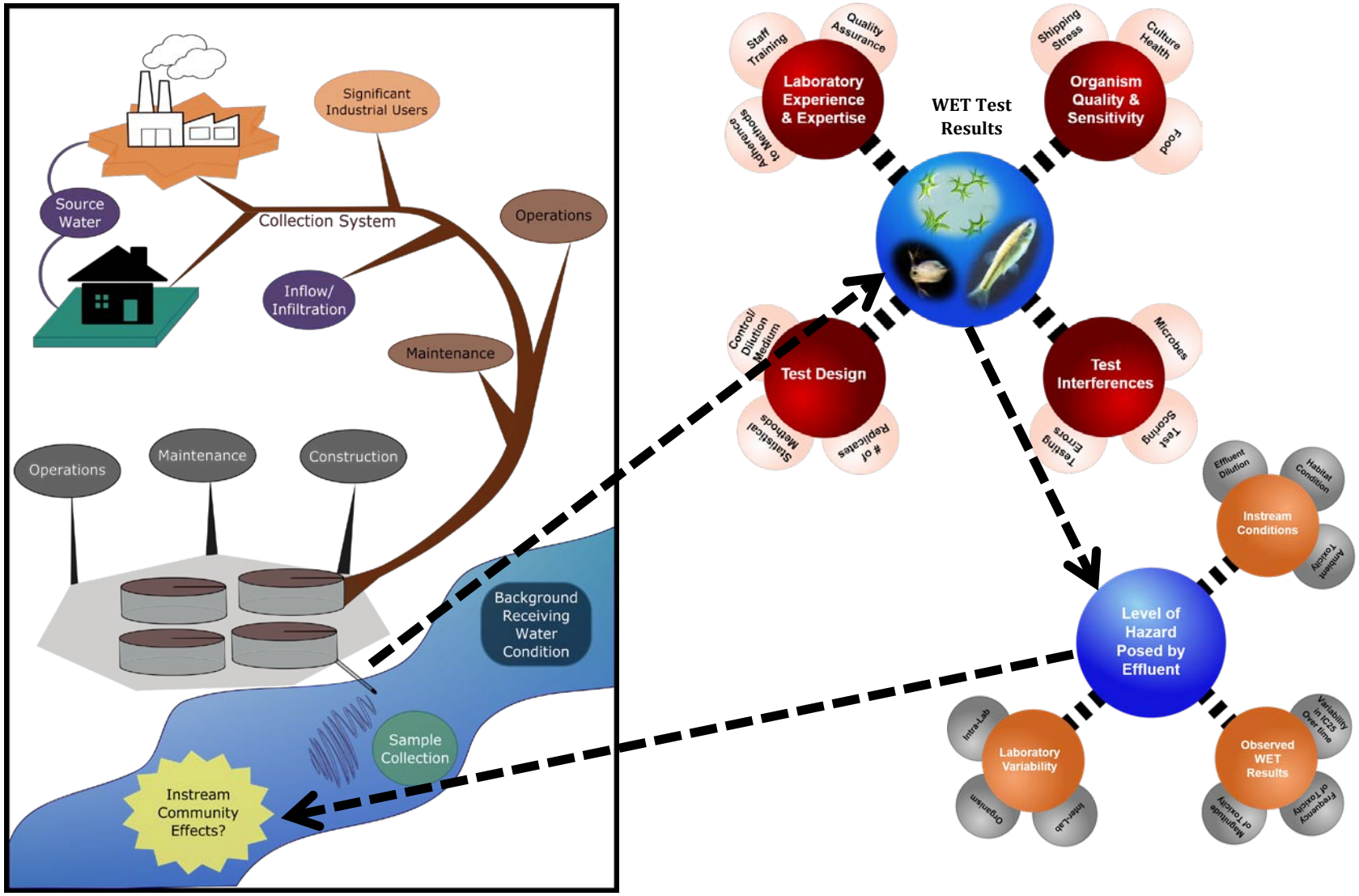

Figure 36. Draft Conceptual Model for Assessing Factors Influencing Chronic Toxicity Test Results and Level of Hazard Posed by Effluent to Instream Aquatic Communities.

This page is intentionally left blank.

Section 8. Summary of Key Findings and Recommendations

Central Valley POTWs are required to conduct periodic three-species (*P. promelas*, *C. dubia*, and *S. capricornutum*) chronic toxicity testing or to use the most sensitive of the three species in chronic toxicity testing to assess the impact that treated effluent may potentially have on the receiving waters and its beneficial uses, including aquatic life. To better understand the nature of the potential issues that surround exceedances of the chronic toxicity trigger, CVCWA conducted this study to characterize the extent to which low-level effects in chronic bioassay tests occur for Central Valley POTWs, identify how exceedances of the chronic toxicity trigger are resolved using the available tools developed and approved by USEPA, evaluate the efficacy of these tools, and develop a conceptual model to better understand numerous variables that can impact the outcome of a chronic toxicity test and the relationship with impairment to instream ecology. From this study, CVCWA may identify, develop, and evaluate additional tools that could be used to better resolve incidents of low-level chronic toxicity in a more effective and cost-efficient way. These additional tools would be proposed to the Central Valley Water Board to supplement existing tools that are currently used by POTWs to investigate and resolve current incidents of identifiable chronic toxicity.

8.1 STUDY KEY FINDINGS

The key findings of this study are discussed below.

8.1.1 Central Valley POTW Chronic Toxicity Characterization

- Central Valley POTWs with a chronic toxicity trigger of 1 TU_c
    - In reviewing chronic toxicity test data from January 2011 to March 2017, Central Valley POTWs with a chronic toxicity trigger of 1 TU_c primarily have exceedances of the chronic toxicity trigger for *C. dubia* reproduction and *S. capricornutum* growth, which are both sub-lethal endpoints. Central Valley POTWs only have isolated incidences of chronic toxicity for *P. promelas* survival and growth and *C. dubia* survival.
    - The majority of the exceedances for Central Valley POTWs with a chronic toxicity trigger of 1 TU_c was 1.3 or 2 TU_c depending on the dilution series that was utilized in the chronic toxicity test. This means that toxicity was observed only in the 100 percent effluent, but not observed in subsequent dilutions during toxicity testing.
    - POTWs using ultraviolet light disinfection observe toxicity for *S. capricornutum* growth twice as frequently as POTWs using chlorination disinfection. Nearly all POTWs utilizing ultraviolet light disinfection have experienced an exceedance of the chronic toxicity trigger for *S. capricornutum* during the data period evaluated.
    - A temporal analysis indicates that the total number of exceedances of the toxicity trigger for *C. dubia* reproduction has increased on a year to year basis.
    - Chronic toxicity trigger exceedances do not exhibit a seasonality trend when all of the chronic toxicity data were evaluated as a whole. However, this may not preclude seasonality trends affecting individual POTWs.

- Central Valley POTWs with a chronic toxicity trigger greater than 1 TU_c
    - In reviewing chronic toxicity test data from January 2011 to March 2017, Central Valley POTWs with a chronic toxicity trigger greater than 1 TU_c primarily have exceedances of the chronic toxicity trigger for *C. dubia* reproduction. These Central Valley POTWs have isolated incidences of chronic toxicity for *P. promelas* survival and growth, *C. dubia* survival, and *S. capricornutum* growth.
    - Because there are few POTWs that have a chronic toxicity trigger greater than 1 TU_c and there were limited chronic toxicity data from these facilities, it was difficult to evaluate potential factors (e.g., treatment level, nitrogen treatment, disinfection methodology) that may influence the outcome of chronic toxicity testing.
- Accelerated testing analysis
    - Based on the available data set, forty percent of accelerated testing conducted by POTWs lead to a TRE.
    - Based on the available data set, fourteen percent of accelerated testing conducted by POTWs did not indicate further chronic toxicity, which allowed POTWs to return to routine monitoring requirements.
    - Because of limitations to the accelerated testing data set, follow-up study of additional accelerated testing can be conducted to improve the data set and refine the understanding of the outcomes of accelerated testing.

8.1.2 Evaluate the Efficacy of TREs and TIEs in Resolving Indications of Effluent Toxicity

- Toxicity Reduction Evaluation Analysis
    - For the majority of TREs that were reviewed, the TREs were resolved through a facility and/or operations review. Accelerated testing for the majority of these incidents were triggered by a relative percent difference of less than 50 percent.
    - Nearly one-quarter of the studies were eventually concluded without identifying the cause or likely cause of the toxicity. In these cases, the likely reason as to why these studies were not resolved is due to lack of persistence in toxicity.
    - TIE testing was conducted as part of 12 TREs, but in only 2 cases were TIE testing effective in identifying the cause of toxicity.
- 'Split-Laboratory' Analysis
    - 'Split-laboratory' studies resulted in a moderate to high degree of agreement with a chronic toxicity triggers. For the 'split-laboratory' comparisons that were performed, the greatest agreement between laboratories occurred for the *P. promelas* test; the laboratories always agreed for this test, but it is important to note that the sample size (n=4) was quite small for this protocol. The *C. dubia* test had the next highest agreement between (and among) laboratories (73.3 to 82.7 percent), and the lowest agreement between laboratories occurred with the *S. capricornutum* test (65 to 77 percent).
    - Typically, there was slightly greater agreement in determining compliance with the trigger between laboratories using a comparison of the IC_{25} as when compared to the NOEC.

8.1.3 Variability of Sub-Lethal Endpoints

- Of the three chronic test species tested by Central Valley POTWs for which inter-laboratory studies have been performed, the *P. promelas* growth exhibited the highest precision, followed by the *S. capricornutum* growth endpoint; the *C. dubia* reproduction endpoint exhibited the lowest precision even though there have been more inter-laboratory studies performed over an estimated 30-year period of time.
- There are a number of POTW, WET testing laboratory, and environmental drivers that can influence the outcome of toxicity tests. When possible, control of the POTW drivers can improve the outcome of the toxicity tests. Similarly, laboratories that have experienced technicians can reduce the influence of some drivers in the laboratory (e.g., organism quality, test interferences, and test design) as can the type of control water selected, thereby minimizing factors that confound the outcome of toxicity tests.
- There are a variety of sources of variability, including numerous sources of intra-test, intra-laboratory, and inter-laboratory variability. All sources of test variability may play a role that can result in different test outcomes between/among laboratories.
- Although the literature provides general sources of intra- and inter-laboratory variability, it would be exceedingly challenging to identify the specific causes of intra- and inter-laboratory variability for the 'split-laboratory' testing evaluated in this study as the compiled 'split-laboratory' testing was not designed investigate the cause of different test outcomes.

8.1.4 Relationship Between Toxicity Testing and Aquatic Ecosystem Impacts

- A literature review performed for this study indicated that it does not appear that WET test results are reliable predictors of effects or lack of effects in the receiving water environment (Chapman 2000, Diamond 2000, Diamond et al. 2008). Specifically, intermittent and low-level toxicity, as measured by sub-lethal endpoints, does not appear to be a reliable predictor of receiving water impairments. Some studies have shown a qualitative correlation between *P. promelas* and *C. dubia* ambient toxicity tests and instream biological condition, but there is considerable debate as to whether these studies are representative of effluents and their receiving waters in general. Since it is particularly challenging to identify levels of sub-lethal endpoint toxicity that do or not correlate with receiving water effects, only general conclusions can be made.
- Ambient water toxicity testing better represents biological condition of the water body than effluent testing, and higher magnitudes of ambient toxicity are better correlated with biological effects (de Vlaming and Norberg-King 1999, Dickson 1992).
- WET testing is better representative of instream biological effects when dilution is considered (de Vlaming and Norberg-King 1999, Diamond et al. 2000, City of Woodland).
- Higher frequencies and magnitude of WET toxicity are generally better correlated with biological effects in a water body (Diamond et al. 2000; Diamond et al. 2008).
- There is no consensus on which WET test species provide the best predictions of biological condition in the receiving water. Different studies have reached different conclusions on this matter (Diamond and Daley 2000, Diamond et al. 2008).

- Because biological responses measured in WET tests are considered less reliable near test detection limits (de Vlaming and Norberg-King 1999), predictions of biological effects in a water body based on WET testing will be improved when laboratory performance and data quality for freshwater chronic WET tests is evaluated with measurement quality objectives that include the use of 'split-laboratory' test evaluation, blind positive control testing, blind negative control testing, and reference toxicants (Diamond et al. 2008).

8.2 STUDY RECOMMENDATIONS

Based on the key findings of this study summarized above, the following recommendations are made to CVCWA for Phase II or subsequent phases of the Toxicity Special Study:

- Conduct Phase II of the Toxicity Special Study, which will further evaluate whether low-level toxicity equates to adverse effects to receiving water aquatic life and beneficial uses, identify and evaluate additional tools that could be utilized by POTWs to investigate low-level indications of toxicity, finalize the conceptual model, and use the technical information compiled and evaluated to refine, expand, and strengthen the toxicity testing process applied to POTWs through NPDES permits. These additional tools can provide an alternative for POTWs to resolve low-level toxicity using methods other than TIE testing if the situation warrants such an approach.
- Refine the accelerated testing data set (e.g., conduct follow-up investigation of the 43 percent of routine chronic toxicity tests that resulted in an indication of toxicity, but toxicity test data during accelerated testing were not available for this study) to better understand the frequency in which Central Valley POTWs conduct TREs or return to routine chronic toxicity monitoring after completion of accelerated testing without a second exceedance of the chronic toxicity trigger. Follow-up investigation can be conducted for the 43 percent of routine chronic toxicity
- Recommend POTWs consider using a third laboratory when 'split-laboratory' testing results in different conclusions to resolve the toxicity and determine compliance through a weight-of-evidence approach.
- Recommend further study into potential causes and correlation of increased chronic toxicity observed for *S. capricornutum* for POTWs using ultraviolet light disinfection in comparison to chlorination-based disinfection.
- Encourage the formation of the SETAC Issues Group and the continuation of SCCWRP-led studies related to addressing and improving inter-laboratory precision for *C. dubia* testing.
- Recommend that POTWs currently conducting a TRE, determine whether the receiving water is currently being impacted by low-level toxicity. In determining whether this would be a useful or beneficial exercise, the criteria at the end of **Section 6** can be reviewed to determine the likelihood that test results will correlate with impacts in the receiving water. It can be reasonably assumed, though not guaranteed, from the literature review conducted that the more criteria that are satisfied, the greater the hazard posed by the effluent to instream biological condition and the more likely that WET results will correlate with receiving water biological condition.

Section 9. References

Central Valley Regional Water Quality Control Board. *Waste Discharge Requirements for Municipal Wastewater Dischargers that Meet Objectives/Criteria at the Point of Discharge to Surface Water* (NPDES No. CAG585001, Order No. R5-2017-0085). August 2017.

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

Clark, S.L. S. Ogle, and J. Cotsifas. 2011. Control treatment water selection can affect the findings of toxicity for *Selenastrum capricornutum* toxicity tests: a source of false positives? 32nd Annual Meeting of the Society of Toxicology and Chemistry North America, Boston, MA.

Cooney, J.D., G.M. DeGraeve, E.L. Moore, B.J. Lenoble, T.L. Pollock, and G.J. Smith. 1992. Effects of environmental and experimental design factors on culturing and toxicity testing of *Ceriodpahnia dubia*. *Environ. Toxicol. Chem.* 11(6): 839-850.

de Vlaming, V., T. J. Norberg-King. A Review Of Single Species Toxicity Tests: Are The Tests Reliable Predictors Of Aquatic Ecosystem Community Responses? U.S. Environmental Protection Agency, Washington, DC, EPA/600/R-97/114, 1999.

DeGrave, G.M., G. Smith, W. Clement, D. McIntyre, and T. Forgette. 1998. *WET Testing Program: Evaluation of Practices and Implementation*. Water Environment Research Foundation. D93002 (94-HHE-1). Alexandria, VA.

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

Diamond, J., J. Stribling, M. Bowersox and H. Latimer. 2008. Evaluation of effluent toxicity as an indicator of aquatic life condition in effluent-dominated streams: a pilot study. Integrated environmental assessment and management 4: 456-470.

Dickson, K.L., W.T. Waller, J.H. Kennedy and L.P. Ammann. 1992. Assessing the Relationship Between Ambient Toxicity And Instream Biological Response. Environ. Toxicol. Chem. 11: 1307-1322.

Environment Canada. 2007. *Biological test method: test and reproduction and survival using the cladoceran Ceriodaphnia dubia*. EPS 1/RM/21, 2nd edition. Method Development and Application Section, Environmental Technology Centre, Ottawa.

Groethe, D.R., K.L. Dickson, and D.K. Reed-Judkins, eds. 1996. Whole Effluent Toxicity Testing; An Evaluation of Methods and Prediction of Receiving System Impacts. Pensacola, FL: SETAC Press.

Marcus, M.D., L.L. McDonald. 1992. Evaluating the Statistical Basis for Relating Receiving Impacts to Effluent and Ambient Toxicities. Environ. Toxicol. Chem. 11: 1389-1402.

Pacholski, L. P. Chapman, A. Hood, and M. Peters. 2017. A cautionary note: *Ceriodaphnia dubia* inter-laboratory test variability. *Bull Environ Contam Toxicol.* 98: 53-57.

Parkhurst, B.R., M.D. Marcus, and L.E. Noel. 1990. Review of the Results of EPA's Complex Effluent Toxicity Testing Program. Utility Water Act Group.

Pennak, R.W. 1978. Freshwater invertebrates of the United States. Protozoa to Mollusca. Third Edition. John Wiley & Sons, Co., New York, NY.

Schiff, K. and D. Greenstein. 2016. *Stormwater Monitoring Coalition: Toxicity Testing Laboratory Guidance Document, Technical Report 956.* Southern California Coastal Water Research Program.

USEPA. 1991. *Technical support document for water quality-based toxics control. Office of Water Enforcement and Permits and Office of Water Regulations and Standards*. US Environmental Protection Agency, Washington, DC, 20460. EPA/505/2-90/001.

USEPA. 1991. Technical Support Document For Water Quality-based Toxics Control. EPA/505/2-90-001. Office of Water. Washington DC. March 1991.

USEPA. 1992. Toxicity Identification Evaluation: Characterization of Chronically Toxic Effluent, Phase I. Office of Research and Development. Washington DC. May 1992.

USEPA 1996. Marine Toxicity Identification Evaluation (TIE) Phase I Guidance Document. EPA/600/R-96/054. September 1996.

USEPA. 1999. Toxicity Reduction Evaluation Guidance for Municipal Wastewater Treatment Plants. EPA/833B-99/002. Office of Wastewater Management. Washington DC. August 1999.

USEPA. 2000. Methods for measuring toxicity and bioaccumulation of sediment-associated contaminants with freshwater invertebrates, 2nd edition. EPA 600/R-99/064. EPA Office of Research and Development (Duluth, MN) and Office of Science and Technology (Washington, D.C.).

USEPA. 2000. *Method Guidance and Recommendations for Whole Effluent Toxicity (WET) Testing (40 CFR Part 136).* Office of Water, U.S. Environmental Protection Agency, Washington, D.C. 20460. EPA-821-B-00-004.

USEPA. 2000. *Understanding and accounting for method variability in whole effluent toxicity applications under the national pollutant discharge elimination system program.* Office of Wastewater Management, U.S. Environmental Protection Agency, Washington, D.C. 20460. EPA/833/R-00/003.

USEPA. 2001. Clarifications Regarding Toxicity Reduction and Identification Evaluations in the National Pollutant Discharge Eliminate System Program. EPA/600/6-91/005F. Office of Wastewater Management. Washington DC. March 27, 2001.

USEPA. 2001. *Final report: interlaboratory variability study of EPA short-term chronic and acute whole effluent toxicity test methods, Vol. 1*. Office of Water, U.S. Environmental Protection Agency, Washington, D.C. 20460. EPA/821/B-01/004.

USEPA. 2002. *Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms, 4th Ed*. Office of Wastewater Management, Washington, D.C. EPA-821-R-02-013.

USEPA. 2007. Sediment Toxicity Identification Manual (TIE): Phases I, II, and III Guidance Document. Office of Research and Development. Washington, D.C. September 2007. EPA/600/R-07/080.

USEPA. 2010. National Pollutant Elimination System Test of Significant Toxicity Implementation Document. Office of Wastewater Management, Washington, D.C. June 2010. EPA/833/R-10/003.

Warren-Hicks, W.J., B.R. Parkhurst. D. Moore, and S. Teed. 1999. *Whole Effluent Toxicity Test Methods: Accounting for Variance*. Water Environment Research Foundation. D83001, Pt. 1&2 (95-PQL-1). Alexandria, VA.

This page is intentionally left blank.

APPENDIX A

Toxicity Report Review Checklist

Review of Laboratory Reports for Whole Effluent Toxicity (WET) Testing

Test review is an important part of a municipalities overall quality assurance program, and is necessary to ensure that all WET test results are reported accurately. The United States Environmental Protection Agency's (USEPA's) *Short-term Methods for Estimating Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms, Fourth Edition* (EPA/821-R-02-013, October 2002) details the principal components of a WET test review in support of a comprehensive quality assurance program, the essential elements of which are detailed in this section and the associated report review checklist.

A comprehensive test review includes the evaluation of the five principal components listed below:

1. Sampling and Sample Handling
2. Test Acceptability Criteria (TAC)
3. Test Conditions
4. Statistical Methods
5. Quality Control

The attached checklist includes the key elements (i.e., method requirements) associated with each of these five principal components. Each element is explicitly detailed as a requirement in USEPA chronic WET testing manual, and thus is similarly a requirement for maintenance of laboratory accreditation. Each element is described in the WET testing manual as "required" or "recommended." Where an element is specified as "required," the method requirement must be met, otherwise the associated test is considered *invalid* for use in determining compliance with a National Pollutant Discharge Elimination System (NPDES) permit's toxicity requirements. When an element is specified as "recommended," the method requirement *should* be met; however, the degree or magnitude of departure from the requirement must be considered in determining whether the associated test result should be invalidated. Ultimately, when a recommended method requirement is exceeded, best professional judgement will need to be employed in the final decision to accept or reject an associated test result.

The attached WET Test Review Worksheet and supporting detailed checklist are provided as tools to assist chronic WET test report reviewers. The WET Test Review Worksheet lists method requirements that the reviewers should verify as being achieved, or if a deviation was recorded, appropriately qualified. The worksheet is provided with space for reviewers to record their notes and/or comments, and can be filed with the reviewed report as part of the reviewer's routine record keeping.

The supporting detailed checklist is organized by principal component, with each essential element and associated test method requirement listed. Where a method requirement is species-specific, additional species-specific guidance is provided. Where appropriate, additional suggestions are listed, with the suggested practices detailing additional guidance for municipalities to consider when reviewing and/or preparing for chronic WET testing.

The worksheet and supporting detailed checklist provide a summary of the principal components and essential elements for an independent review of chronic WET test data. These documents

cover those items most often related to chronic WET test data qualification and/or invalidation. Each chronic WET test method also includes recommended and required test conditions that rarely are of issue, such as light quality, light intensity, photoperiod, feeding, cleaning, etc. While not included in the worksheet or supporting detailed checklist, these additional test conditions can be found in the individual "Summary of Test Conditions" table of each WET test method of the USEPA method protocol (EPA 821-R-02-013); see Attachment A for the Summary of Test Conditions for each species.

A comprehensive test review will require obtaining all available information pertaining to the five principal test review components and associated elements listed in the following checklist. A laboratory report that does not provide the "Method Requirement" information contained in the following checklist will prohibit the municipality (and regulatory authority) from performing an independent validation of the reported WET test results. However, in order for a laboratory to maintain laboratory accreditation (Environmental Laboratory Accreditation Program [ELAP] and/or National Environmental Laboratory Accreditation Program [NELAP]), the laboratory must collect, document, and maintain all information listed in the following checklist. Therefore, if found absent in the laboratory's standard WET test report, the information can be requested from the laboratory.

Attachment A

Chronic Freshwater Toxicity Test Standard Test Condition Sheets

TABLE 1. SUMMARY OF TEST CONDITIONS AND TEST ACCEPTABILITY CRITERIA FOR FATHEAD MINNOW, *PIMEPHALES PROMELAS*, LARVAL SURVIVAL AND GROWTH TOXICITY TESTS WITH EFFLUENTS AND RECEIVING WATERS (TEST METHOD 1000.0)[1]

1.	Test type:	Static renewal (required)
2.	Temperature (°C):	25 ± 1°C (recommended) Test temperatures must not deviate (i.e., maximum minus minimum temperature) by more than 3°C during the test (required)
3.	Light quality:	Ambient laboratory illumination (recommended)
4.	Light intensity:	10-20 μE/m^2/s (50-100 ft-c)(ambient laboratory levels) (recommended)
5.	Photoperiod:	16 h light, 8 h darkness (recommended)
6.	Test chamber size:	500 mL (recommended minimum)
7.	Test solution volume:	250 mL (recommended minimum)
8.	Renewal of test solutions:	Daily (required)
9.	Age of test organisms:	Newly hatched larvae less than 24 h old. If shipped, not more than 48 h old, 24 h range in age (required)
10.	No. larvae per test chamber:	10 (recommended)
11.	No. replicate chambers per concentration:	4 (required minimum)
12.	No. larvae per concentration:	40 (required minimum)
13.	Source of food:	Newly hatched *Artemia* nauplii (less than 24 h old) (required)
14.	Feeding regime:	On days 0-6, feed 0.1 g newly hatched (less than 24-h old) brine shrimp nauplii three times daily at 4-h intervals or, as a minimum, 0.15 g twice daily at 6-h intervals (at the beginning of the work day prior to renewal, and at the end of the work day following renewal). Sufficient nauplii are added to provide an excess. (recommended)

[1] For the purposes of reviewing WET test data submitted under NPDES permits, each test condition listed above is identified as required or recommended (see Subsection 10.2 for more information on test review). Additional requirements may be provided in individual permits, such as specifying a given test condition where several options are given in the method.

TABLE 1. SUMMARY OF TEST CONDITIONS AND TEST ACCEPTABILITY CRITERIA FOR FATHEAD MINNOW, *PIMEPHALES PROMELAS*, LARVAL SURVIVAL AND GROWTH TOXICITY TESTS WITH EFFLUENTS AND RECEIVING WATERS (TEST METHOD 1000.0) (CONTINUED)

15.	Cleaning:	Siphon daily, immediately before test solution renewal (required)
16.	Aeration:	None, unless DO concentration falls below 4.0 mg/L. Rate should not exceed 100 bubbles/minimum (recommended)
17.	Dilution water:	Uncontaminated source of receiving or other natural water, synthetic water prepared using MILLIPORE MILLI-Q® or equivalent deionized water and reagent grade chemicals, or DMW (see Section 7, Dilution Water) (available options)
18.	Test concentrations:	Effluents: 5 and a control (required minimum) Receiving Water: 100% receiving water (or minimum of 5) and a control (recommended)
19.	Dilution factor:	Effluents: ≥ 0.5 (recommended) Receiving waters: None or ≥ 0.5 (recommended)
20.	Test duration:	7 days (required)
21.	Endpoints:	Survival and growth (weight) (required)
22.	Test acceptability criteria:	80% or greater survival in controls; average dry weight per surviving organism in control chambers equals or exceeds 0.25 mg (required)
23.	Sampling requirements:	For on-site tests, samples collected daily, and used within 24 h of the time they are removed from the sampling device; For off-site tests, a minimum of three samples (e.g., collected on days one, three and five) with a maximum holding time of 36 h before first use (see Section 8, Effluent and Receiving Water Sampling, Sample Handling, and Sample Preparation for Toxicity Tests, Subsection 8.5.4) (required)
24.	Sample volume required:	2.5 L/day (recommended)

TABLE 3. SUMMARY OF TEST CONDITIONS AND TEST ACCEPTABILITY CRITERIA FOR DAPHNID, *CERIODAPHNIA DUBIA*, SURVIVAL AND REPRODUCTION TOXICITY TESTS WITH EFFLUENTS AND RECEIVING WATERS (TEST METHOD 1002.0)[1]

1.	Test type:	Static renewal (required)
2.	Temperature (°C):	25 ± 1°C (recommended) Test temperatures should not deviate (i.e., maximum minus minimum temperature) by more than 3°C during the test (required)
3.	Light quality:	Ambient laboratory illumination (recommended)
4.	Light intensity:	10-20 μE/m^2/s, or 50-100 ft-c (ambient laboratory levels) (recommended)
5.	Photoperiod:	16 h light, 8 h dark (recommended)
6.	Test chamber size:	30 mL (recommended minimum)
7.	Test solution volume:	15 mL (recommended minimum)
8.	Renewal of test solutions:	Daily (required)
9.	Age of test organisms:	Less than 24 h; and all released within a 8-h period (required)
10.	No. neonates per test chamber:	1 Assigned using blocking by known parentage (Subsection 13.10.2.4) (required)
11.	No. replicate test chambers per concentration:	10 (required minimum)
12.	No. neonates per test concentration:	10 (required minimum)
13.	Feeding regime:	Feed 0.1 mL each of YCT and algal suspension per test chamber daily (recommended)
14.	Cleaning:	Use freshly cleaned glass beakers or new plastic cups daily (recommended)
15.	Aeration:	None (recommended)
16.	Dilution water:	Uncontaminated source of receiving or other natural water, synthetic water prepared using MILLIPORE MILLI-Q® or equivalent deionized water and reagent grade chemicals or DMW (see Section 7, Dilution Water) (available options)

[1] For the purposes of reviewing WET test data submitted under NPDES permits, each test condition listed above is identified as required or recommended (see Subsection 10.2 for more information on test review). Additional requirements may be provided in individual permits, such as specifying a given test condition where several options are given in the method.

TABLE 3. SUMMARY OF TEST CONDITIONS AND TEST ACCEPTABILITY CRITERIA FOR DAPHNID, *CERIODAPHNIA DUBIA*, SURVIVAL AND REPRODUCTION TOXICITY TESTS WITH EFFLUENTS AND RECEIVING WATERS (TEST METHOD 1002.0) (CONTINUED)

17.	Test concentrations:	Effluents: 5 and a control (required minimum) Receiving Water: 100% receiving water (or minimum of 5) and a control (recommended)
18.	Dilution factor:	Effluents: ≥ 0.5 (recommended) Receiving Waters: None or ≥ 0.5 (recommended)
19.	Test duration:	Until 60% or more of surviving control females have three broods (maximum test duration 8 days) (required)
20.	Endpoints:	Survival and reproduction (required)
21.	Test acceptability criteria:	80% or greater survival of all control organisms and an average of 15 or more young per surviving female in the control solutions. 60% of surviving control females must produce three broods (required)
22.	Sampling requirements:	For on-site tests, samples collected daily and used within 24 h of the time they are removed from the sampling device. For off-site tests, a minimum of three samples (e.g., collected on days one, three, and five) with a maximum holding time of 36 h before first use (see Section 8, Effluent and Receiving Water Sampling, Sample Handling, and Sample Preparation for Toxicity Tests, Subsection 8.5.4) (required)
23.	Sample volume required:	1 L/day (recommended)

TABLE 3. SUMMARY OF TEST CONDITIONS AND TEST ACCEPTABILITY CRITERIA FOR GREEN ALGA, *SELENASTRUM CAPRICORNUTUM,* GROWTH TOXICITY TESTS WITH EFFLUENTS AND RECEIVING WATERS (TEST METHOD 1003.0)[1]

1.	Test type:	Static non-renewal (required)
2.	Temperature:	25 ± 1 °C (recommended) Test temperatures must not deviate (i.e., maximum minus minimum temperature) by more than 3°C during the test (required)
3.	Light quality:	"Cool white" fluorescent lighting (recommended)
4.	Light intensity:	86 ± 8.6 μE/m^2/s (400 ± 40 ft-c or 4306 lux) (recommended)
5.	Photoperiod:	Continuous illumination (required)
6.	Test chamber size:	125 mL or 250 mL (recommended)
7.	Test solution volume:	50 mL or 100 mL[2] (recommended)
8.	Renewal of test solutions:	None (required)
9.	Age of test organisms:	4 to 7 days (required)
10.	Initial cell density in test chambers:	10,000 cells/mL (recommended)
11.	No. replicate chambers per concentration:	4 (required minimum)
12.	Shaking rate:	100 cpm continuous, or twice daily by hand (recommended)
13.	Aeration:	None (recommended)
14.	Dilution water:	Algal stock culture medium, enriched uncontaminated source of receiving or other natural water, synthetic water prepared using MILLIPORE MILLI-Q® or equivalent deionized water and reagent grade chemicals, or DMW (see Section 7, Dilution Water) (available options)

[1] For the purposes of reviewing WET test data submitted under NPDES permits, each test condition listed above is identified as required or recommended (see Subsection 10.2 for more information on test review). Additional requirements may be provided in individual permits, such as specifying a given test condition where several options are given in the method.

[2] For tests not continuously shaken use 25 mL in 125 mL flasks and 50 mL in 250 mL flasks.

TABLE 3. SUMMARY OF TEST CONDITIONS AND TEST ACCEPTABILITY CRITERIA FOR GREEN ALGA, *SELENASTRUM CAPRICORNUTUM,* GROWTH TOXICITY TESTS WITH EFFLUENTS AND RECEIVING WATERS (TEST METHOD 1003.0) (CONTINUED)

15.	Test concentrations:	Effluents: 5 and a control (required minimum) Receiving Water: 100% receiving water (or minimum of 5) and a control (recommended)
16.	Test dilution factor:	Effluents: ≥ 0.5 (recommended) Receiving Waters: None or ≥ 0.5 (recommended)
17.	Test duration:	96 h (required)
18.	Endpoint:	Growth (cell counts, chlorophyll fluorescence, absorbance, or biomass) (required)
19.	Test acceptability criteria:[3]	Mean cell density of at least 1 X 10^6 cells/mL in the controls; and variability (CV%) among control replicates less than or equal to 20% (required)
20.	Sampling requirements:	For on-site tests, one sample collected at test initiation, and used within 24 h of the time it is removed from the sampling device. For off-site tests, holding time must not exceed 36 h before first use (see Section 8, Effluent and Receiving Water Sampling, Sample Handling, and Sample Preparation for Toxicity Tests, Subsection 8.5.4) (required)
21.	Sample volume required:	1 or 2 L depending on test volume (recommended)

[3] If the test is conducted under non-NPDES applications (i.e., data are not submitted under NPDES permits) and used without EDTA in the nutrient stock solution, the test acceptability criteria are a mean cell density of at least 2 X 10^5 cells/mL in the controls, and variability (CV%) among control replicates less than or equal to 20%.

Attachment B

Chronic Toxicity Report Review Checklists

WET TEST REVIEW WORKSHEET

	Element	Method Requirement	Report Reviewer Comment
☐	Sample Transport	Sample should be delivered chilled to 0-6°C during or immediately after collection (recommended).	
☐	Sample Holding Time	Sample holding time shall not exceed 36 hours from collection to first use (required).	
☐	Sample Usage	In static renewal, each sample may be used to prepare test solutions for a maximum of 72 hours after first use (required).	
☐	Renewal Samples	For chronic *C. dubia* and *P. promelas* tests, a minimum of two renewal samples are to be collected (required).	
☐	Test Acceptability Criteria	Species specific TAC must be achieved for test to be considered valid (required).	
☐	Temperature	Test temperatures should be 25 ± 1°C (recommended).	
☐	Temperature	The minimum minus maximum temperature for the test must not deviate by more than 3°C (required).	
☐	Dissolved Oxygen	Samples should be gently aerated if sample is received with dissolved oxygen below 4 mg/L (recommended).	
☐	Dissolved Oxygen	Sample dissolved oxygen should be near 100% saturation prior to use (recommended).	

WET TEST REVIEW WORKSHEET

	Element	Method Requirement	Report Reviewer Comment
☐	Dissolved Oxygen	Test solutions are not aerated unless dissolved oxygen in test solutions falls below 4 mg/L (recommended).	
☐	Age of Organisms	Tests are to be initiated with organisms of known age and within the age range specified for the test (required).	
☐	Replication	Tests are required to be initiated with the minimum required number of replicates and organisms per replicate specified in each test method (required).	
☐	Statistical Flow Chart	Method specific statistical flow chart should be followed (recommended).	
☐	Concentration-Response Relationship	The concentration-response relationship generated for a multi-concentration test must be reviewed to evaluate the validity of test results (required).	
☐	Reference Toxicant Testing	The associated concurrent reference toxicant test should be reviewed and test results falling outside of laboratory specific control limits explained (recommended).	
☐	Test Variability (PMSD)	When NPDES permits require sub-lethal hypothesis testing, within-test variability must be reviewed and PMSD variability criteria must be applied (required).	

Sampling and Sample Handling

	Element	Method Requirement	Species-Specific Guidance	Suggested Practice
☐	Sample Transport	▶ Sample should be delivered chilled to 0-6°C during or immediately after collection (recommended).	N/A	▶ Samples delivered the same calendar day as collection need only to show evidence of cooling. To provide evidence of cooling: 1) indicate sample temperature at time of sample collection on COC – a laboratory sample log-in temperature less than the temperature indicated on COC provides evidence of cooling, or 2) a temperature blank at 0-6ºC upon receipt at the laboratory demonstrated adequate ice was used to chill the sample during transport.
☐	Sample Holding Time	▶ Sample holding time shall not exceed 36 hours from collection to first use (required). o Requirement applies to initial and renewal samples. o A variance to this holding time may be obtained from USEPA, but under no circumstance should the time between collection and first use be greater than 72 hours. o If shipping problems (e.g., unsuccessful Saturday delivery) are encountered with renewal samples after a test has been initiated, the Central Valley Water Board may allow the continued use of the most recently used sample for test renewal.	N/A	▶ Sample collection time indicated on COC should represent the time of day a grab sample was collected or the time of day the last aliquot of a composite sample was drawn. ▶ For samples being shipped to off-site testing laboratories, sample collection should occur in the morning, the day of shipment, after 6 am. This will ensure the laboratory has adequate time to initiate the tests within hold time.
☐	Sample Usage	▶ In static renewal tests (i.e., chronic *C. dubia* and *P. promelas*), each sample may be used to prepare test solutions for solution renewal for a maximum of 72 hours after first use (required).	N/A	▶ Samples should be stored at 0-6˚C with minimum headspace.
☐	Renewal Samples	▶ For chronic *C. dubia* and *P. promelas* tests, a minimum of two renewal samples are to be collected (required).	▶ For the chronic S*. capricornutum* test, only a single sample need be collected.	▶ For chronic *C. dubia* and *P. promelas* tests, an every-other-day sample collection schedule (i.e., M, W, F) ensures the method solution renewal and sample usage requirement is achieved. Moreover, an every-other-day sample collection schedule ensures the duration of organism exposure to any single sample is as balanced (i.e., the test organisms are exposed to the samples for a similar duration of time thus minimizing sample exposure bias). ▶ Consecutive sampling schedules (e.g., M, Tu, W) will not achieve method required sample usage requirements for chronic *C. dubia* and *P. promelas* test, and thus will invalidate the test for NPDES compliance monitoring purposes.

Test Acceptability Criteria (TAC)

	Element	Method Requirement	Species-Specific Guidance	Suggested Practice
☐	Test Acceptability Criteria	► Species-specific TAC must be achieved for test to be considered valid (required).	<u>*S. capricornutum*</u> ► ≥1x10^{6} average cell density in control ► Control coefficient of variation ≤0.2 (i.e., ≤20%) <u>*C. dubia*</u> ► ≥80% survival in control ► ≥15 average young per surviving female in control ► ≥60% of surviving control females produce 3 broods, to be achieved in a maximum of 8 days. <u>*P. promelas*</u> ► ≥80% survival in control ► Average dry weight of surviving organisms ≥0.25 mg	► In most cases, if TAC are not achieved for a test, that test must be repeated with fresh samples. See NPDES permit Monitoring and Reporting Program (MRP) for details. ► TAC are method protocol required minimum standards of test performance that should be regularly surpassed. Species-specific guidance is provided below. <u>*S. capricornutum*</u> ► Plating of algal cells on the bottom of the test chambers has been reported for a number of POTWs. As this plating reduces the cell counts in the suspended water, laboratories should re-suspend plated algae prior to measurement or enumeration. ► This test requires amending the laboratory control water and effluent samples with nutrients to promote algal growth, which is then compared to the TAC. As many control water (e.g., well, reconstituted, 80:20) already have base nutrients that can further promote algal growth, pure water (e.g., Type I/distilled water/deionized) is preferred as the laboratory control base media to reduce the possibility of 'false positives'. <u>*C. dubia*</u> ► The minimum TAC requires at least 15 average young per surviving female. The method protocol states the following regarding the manner in which *C. dubia* brood production typically transpires over the course of the test: "Three or four broods are usually obtained in the controls in a 7-day test conducted at 25 ± 1°C. A brood is a group of offspring released from the female over a short period of time when the carapace is discarded during molting. In the controls, the first brood of two-to-five young is usually released on the third or fourth day of the test. Successive broods are released every 30 to 36 h thereafter. The second and third broods usually consist of eight to 20 young each. The total number of young produced by a healthy control organism in three broods often exceeds 30 per female" (EPA 821-R-02-013, §13.10.6.2.10). Control reproduction in and around the minimum of 15 required to meet TAC may, but does not necessarily, indicate unhealthy or poor quality organisms. <u>All organisms</u> ► Abrupt changes in control growth or reproduction (i.e., exceedingly low or exceedingly high) should be evaluated to determine if it is a deviation from the laboratory's typical performance and thus indicative of a possibly anomalous result deserving qualification. Control growth or reproduction should be reviewed in the context of the laboratory's typical performance using control charts. ► Evaluation of laboratory control charts for mean control performance and related control %CV will provide a measure of the laboratories performance over time (i.e., one year). POTWs should consider requesting such control charts to address long-term laboratory performance.

Test Conditions

	Element	Method Requirement	Species-Specific Guidance	Suggested Practice
☐	Temperature	► Test temperatures should be 25 ±1°C (recommended). ► The maximum minus minimum temperature for the test must not deviate by more than 3°C (required).	N/A	
☐	Dissolved Oxygen	► Samples should be gently aerated if sample is received with dissolved oxygen below 4 mg/L (recommended). ► Sample dissolved oxygen should be near 100% saturation prior to use (recommended). o Warming samples slowly in open containers to test temperature or gentle aeration may be used. ► Test solutions are not aerated unless dissolved oxygen in test solutions falls below 4 mg/L (recommended). o If aeration of test solutions is necessary, all treatments including controls must be aerated.	N/A	► If aeration is necessary during testing, samples and/or test solutions should be gently aerated at a rate of not greater than 100 bubbles per minute and in a manner that does not cause turbulence and undue stress on the test organism. ► Aeration of chronic *C. dubia* test solutions is generally not practical given the delicate nature of the test organism. Aeration of test solutions prior to use is typically the most effective means of addressing dissolved oxygen issues for this species. ► Aeration of samples or test solutions should be minimized to avoid loss of volatile chemicals, thus aeration to address dissolved oxygen problems is a method of last resort.
☐	Age of Organisms	► Tests are to be initiated with organisms of known age and within the age range specified for the test (required).	*S. capricornutum* ► Cell culture must be 4 to 7 days old at time of initiation. *C. dubia* ► Organisms must be less than 24 hours old and within an 8-hour age range (e.g., 0-8, 9-16, 17-24 hour old) at time of initiation. *P. promelas* ► Organisms must be less than 48 hours old at time of initiation.	*S. capricornutum* ► Laboratory is required to maintain culture logs documenting the age, and must also record the age of the culture used for each test. *C. dubia* ► Laboratory is required to maintain documentation that neonates used for testing came from the third or subsequent brood, and maintain records to support that the neonates were less than 24 hour old and within 8 hours of age. *P. promelas* ► If not cultured by the laboratory performing the testing, the laboratory must maintain records with the hatch date and time from their culture vendor to support that the larvae used for testing were less than 48 hours old at test initiation. ► For fish not cultured by the laboratory, less than 48 hours old fish can only be received on days shipments can be received, usually Tuesday through Saturday.
☐	Replication	► Tests are required to be initiated with the minimum required number of replicates and organisms per replicate specified in each test method (required).	*S. capricornutum* ► A minimum of four replicates per treatment. ► Each replicate should be inoculated at an initial cell density of 10,000 cells/mL. *C. dubia* ► A minimum of ten replicates per treatment. ► Each replicate must be loaded with one organism, using “blocking by known parentage”. *P. promelas* ► A minimum of four replicates per treatment. ► Each replicate must be loaded with a minimum of 10 organisms.	► Test sensitivity and precision generally increases with increasing replication. Where increased sensitivity and precision are desired (e.g., where the Test of Significant Toxicity is used), increased replication should be considered. ► Pathogen related mortality/toxicity in the chronic *P. promelas* test can be mitigated by increasing replication to 20 with each replicate containing a minimum of two organisms.

Statistical Methods

	Element	Method Requirement	Species-Specific Guidance	Suggested Practice
☐	Statistical Flow Chart	► Method-specific statistical flow chart should be followed (recommended).	N/A	► Deviations from the recommended method-specific statistical flow chart is permitted, but must be noted in the laboratory report, including justification for the deviation and demonstration that the statistical model assumptions are achieved. ► Use of alternative statistical models and methods (i.e., probit analysis, Test of Significant Toxicity) should be considered when building a weight-of-evidence approach to evaluating the presence of low-level and/or threshold toxicity. However, such methods require regulatory approval. ► Where anomalous concentration-response curves are observed, best professional judgement should be utilized to determine the appropriateness of the statistical analyses outlined in the method-specific flow chart.

Quality Control

	Element	Method Requirement	Species-Specific Guidance	Suggested Practice
☐	Concentration-Response Relationship	▶ The concentration-response relationship generated for a multi-concentration test (e.g., dilution series) must be reviewed to evaluate the validity of test results (required). o Where testing only employs a single treatment, evaluation of the concentration-response is not possible, and thus not required.	N/A	▶ Even when an allowance to test a single treatment is permitted, testing with a dilution series should be considered to allow for additional quality control evaluation of the resultant concentration-response relationship. ▶ The USEPA guidance document "Method Guidance and Recommendations for Whole Effluent Toxicity Testing" (EPA-821-B-00-004) provides example anomalous concentration-response relationships and recommended investigative and corrective actions that should be taken if they are observed in test results, including whether a test result should be determined reliable and reported, a test result should be determined anomalous and explained, or a test result should be determined inconclusive and the test repeated. ▶ Where anomalous concentration-response curves are observed, best professional judgement should be utilized to determine the appropriateness of the statistical analyses outlined in the method-specific flow chart.
☐	Reference Toxicant Testing	▶ The associated concurrent reference toxicant test is reviewed and test results falling outside of laboratory specific control limits explained (recommended). o Concurrent reference toxicant tests are currently required in Central Valley NPDES permits	▶ Specific recommended control limits will vary by species and laboratory, but should be presented as the mean relevant effect threshold (e.g., EC_{50}, IC_{50}) ±2 standard deviations for the last 20 tests performed by the laboratory.	▶ A reference toxicant test result outside laboratory control limits will not necessarily invalidate an associated effluent test result. The direction and magnitude of the control limit exceedance should be taken into account as to whether an associated effluent test result should be accepted or rejected. For example, where a reference toxicant test result that shows greater than typical sensitivity concurrent with no observable toxicity in the effluent test, the effluent test result can be accepted as no toxicity was observed despite the atypically higher sensitivity. In the inverse, where atypically higher sensitivity is associated with an observation of toxicity in a concurrent effluent test, the effluent test can be appropriately qualified, and the test result judiciously interpreted.
☐	Test Variability (Percent Minimum Significant Difference)	▶ When NPDES permits require sub-lethal hypothesis testing (i.e., growth and reproduction NOECs), within-test variability must be reviewed and variability criteria (e.g., method specific PMSD limits) must be applied (required). o If test PMSD exceeds upper bound and effluent is not toxic, test must be repeated. o If test PMSD exceeds upper bound and effluent is toxic, accept test. o If test PMSD is less than lower bound and effluent is not toxic, accept test. o If test PMSD is less than lower bound and effluent is toxic but relative percent difference between control and effluent is less than the lower PMSD bound, test is accepted and effluent is determined not toxic.	*S. capricornutum* ▶ PMSD upper bound: 29% ▶ PMSD lower bound: 9.1% *C. dubia* ▶ PMSD upper bound: 47% ▶ PMSD lower bound: 13% *P. promelas* ▶ PMSD upper bound: 30% ▶ PMSD lower bound: 12%	▶ In addition to its required use in NPDES compliance testing, PMSD is an overall good indicator of laboratory performance. Evaluation of laboratory control charts of test specific PMSD should be considered when performing long-term evaluations of laboratory performance.

This page is intentionally left blank.