OC Public Works
Integrity, Accountability, Service, Trust
Shane L. Silsby, Director

July 26, 2017

Mr. David W. Gibson
Executive Officer,
California Regional Water Quality Control Board, San Diego Region
2375 Northside Drive, Suite 100
San Diego, CA 92108

SUBJECT: Orange County MS4 Monitoring and Assessment Program Recommendation - *Ceriodaphnia dubia* Toxicity Testing

Dear Mr. Gibson:

The County of Orange, the Orange County Flood Control District, and the incorporated Cities of Orange County within the San Diego Region (Permittees) are currently implementing the transitional water quality monitoring and assessment program required by Order R9-2013-0001, as amended by R9-2015-0001 and R9-2015-0100 (Order). The Permittees have also developed a long term monitoring and assessment program (MAP) as part of the proposed Water Quality Improvement Plan for the South Orange County watershed management area. An element of both the transitional monitoring program and the proposed WQIP is water column toxicity monitoring for survival and reproduction of *Ceriodaphnia dubia* (*C. dubia*) in accordance with provisions D.1.c.(4)(a)-(f) and D.1.d.(4) for both dry and wet weather receiving water monitoring. Water column toxicity monitoring for *C. dubia* has also been part of the historical long term mass emissions monitoring conducted by the Permittees.

The County, as Principal Permittee, actively participates in the Southern California Stormwater Monitoring Coalition (SMC) on behalf of the Permittees. A recent toxicity intercalibration study funded by the SMC and coordinated by the Southern California Coastal Water Research Project (SCCWRP) identified poor comparability with toxicity testing, particularly between laboratories and within individual laboratories. One of the most concerning revelations was related to the *C. dubia* reproduction test where a blind sample of laboratory dilution water, prepared in accordance with standardized guidance and expected to be nontoxic, was found to be toxic by some of the participating labs. This prompted a review of test methods and development of a laboratory guidance document (see attachment) prior to a second round of testing. The results of the second round of interlaboratory testing also identified toxicity in the dilution water sample and relative low comparability of the laboratories between testing events. A cause was not identified and further investigation of the issue was recommended.

The SMC has proposed a third round of toxicity intercalibration study to identify laboratory quality assurance practices that will improve comparability of the *C. dubia* reproduction test. The focus of this work is to identify test conditions and procedures that will minimize instances of toxicity when

laboratory dilution water is tested by multiple laboratories. Improving the comparability of test results for samples expected to be nontoxic will improve confidence in the use and interpretation of the *C. dubia* reproduction test for effluent and stormwater monitoring. One of the outcomes will be an additional guidance document that describes test procedures and quality assurance steps to improve toxicity data comparability.

A third round of study would again be coordinated by SCCWRP, and it is estimated that it will take 36 months to complete, at a cost of $700K. Approximately 50% of project funding will be used to compensate testing laboratories for participating in the study (see attached concept proposal). Given the current uncertainty over *C. dubia* toxicity results, and the potential implications, a third round of the study is critical and we respectfully request your concurrence of a proposal to modify the MAP by eliminating *C. dubia* survival and reproduction toxicity testing until such time as the comparability issue has been addressed. In exchange, the Permittees would fund this study along with other participating SMC organizations. All other water column toxicity monitoring would continue without modification.

This proposed modification is consistent with provision H.4.e of the Order, which authorizes updating or revising the monitoring and reporting requirements, at the discretion of the San Diego Water Board, based on recommendations from SCCWRP (among other justifications). The Permittees believe that this request is well justified and will promote development and implementation of improved monitoring and assessment programs throughout the Southern California region. An identical request has been submitted to the Executive Officer of the Santa Ana Regional Water Board for consideration.

If you have any questions, require additional information, or would like to discuss further details, please contact James Fortuna at (714) 955-0680.

Sincerely,

Grant Sharp, Manager,
Environmental Monitoring Division

Attachments: 1) Development of Quality Assurance Guidance for the *Ceriodaphnia dubia* Reproduction Test
2) Ceriodaphnia QA Development Concept Proposal

cc: Arne Anselm, (SMC Chair)
Ken Schiff, SCCWRP
Barbara Barry, Santa Ana Water Board
Adam Fischer, Santa Ana Water Board (SMC Rep)
Chad Loflen, San Diego Regional Water Board (SMC Rep)
Laurie Walsh, San Diego Water Board
Erica Ryan, San Diego Water Board
Roger Mitchell, San Diego Water Board

Attachment 1

STORMWATER MONITORING COALITION

TOXICITY TESTING LABORATORY GUIDANCE DOCUMENT

FINAL DRAFT

NOVEMBER 2016

EXECUTIVE SUMMARY

Aquatic toxicity testing has become a standard measurement in stormwater management. Samples collected in the field are brought back to the laboratory, where test organisms are exposed and their response – ranging from lethality to critical life stage development or reproduction success – is measured using very uniform and repeatable methods. Cumulatively, stormwater management agencies in southern California spend nearly $1 million annually conducting toxicity tests.

The Southern California Stormwater Monitoring Coalition (www.SoCalSMC.org) includes 15 regulated and regulatory agencies from Ventura to San Diego, and one of their goals is to combine data sets for making comparisons between watersheds or over time. One challenge to using toxicity testing is that the various SMC member agencies currently utilize different test species and a variety of endpoints. Although standardized methods are used by the multiple contract laboratories who conduct SMC toxicity testing, the method protocols typically have options or interpretations left to the laboratory, potentially leading to different test outcomes. This uncertainty is compounded by concerns about the toxicity test's inherent variability within each laboratory.

As a result of these challenges, the SMC decided to conduct a laboratory intercalibration study to assess comparability. The goal was to identify some key recommended test species and endpoints, quantify intra- and inter-laboratory variability for each test, and make recommendations for how to minimize that variability, where applicable. An Advisory Committee was created to help design, implement, and interpret the intercalibration study, then construct the recommendations in this Guidance Manual.

The recommended test species include two freshwater species (*Ceriodaphnia dubia* 6-8 day chronic survival and reproduction test and *Hyalella azteca* 96-hour acute survival test) and two marine species (*Strongylocentrotus purpuratus* and *Mytilus galloprovincialis* short-term chronic larval development tests) based on commonality to current monitoring requirements and maintaining existing trends, sensitivity to toxicants, and ease of testing/cost, amongst other criteria. Two iterations of laboratory intercalibrations were conducted. Each iteration was comprised of four samples, delivered blind to each laboratory; lab dilution water, lab dilution water spiked with copper, runoff sample created with artificial rainfall, and a duplicate. Comparability was evaluated based on three factors; test acceptability (negative control and reference toxicant response), intra-laboratory precision (duplicate sample response), and inter-laboratory precision (among lab response). Up to 10 laboratories participated including contract labs, municipal monitoring labs, and research labs. All of the laboratories were certified by the State of California for toxicity testing.

After two intercalibration iterations, nearly all laboratories scored comparable (moderate to very high comparability) for three of the four species (four of five endpoints) including both marine species, *Hyalella* (the newest method), and the survival endpoint for *Ceriodaphnia* (Table ES-1) However, approximately half the laboratories scored moderate or better comparability for the *Ceriodaphnia* reproduction test, and these laboratories were not consistent between intercalibration rounds. While intra-laboratory precision was generally comparable for *Ceriodaphnia* reproduction, there was a range of responses among laboratories to each sample, including the lab dilution water. The best inter-laboratory precision for the *Ceriodaphnia* reproduction test was observed for the runoff sample.

Table ES-1. Summary of laboratory comparability scoring for *Ceriodaphnia dubia* (6-8 day) survival and reproduction, *Hyalella* survival, *Strongylocentrotus* embryo development, or *Mytilus* embryo development tests.

Lab	*Ceriodaphnia* Survival		*Ceriodaphnia* Reproduction		*Hyalella* Survival		*Strongylo-centrotus* Development	*Mytilus* Embryo Development
	Round 1	*Round 2*	*Round 1*	*Round 2*	*Round 1*	*Round 2*	*Round 1* [a]	*Round 1* [a]
A	Moderate	High	Very High	Low	Low	High	Moderate	- [b]
B	Very High	High	Moderate	High	Low	High	-	-
C	Low	High	Low	High	Low	Very High	-	-
E	Moderate	-	Moderate	-	-		-	Very High
F	Moderate	High	Moderate	Low	Low	Very High	Moderate	Low
G	High	-	High	-	-	-	-	-
H	Low	-	Low	-	-	-	-	-
I	High	Moderate	High	Low	Moderate	Very High	High	Very High
J	Low	High	Low	Low	High	Very High	Moderate	Moderate

[a] Only tested in Round 1
[b] – indicates sample not tested

Based on these results, all four species can be recommended for future use as part of the SMC monitoring programs. Specific guidance for stormwater testing is given for potentially variable inducing steps including hardness of dilution water, feeding, sample handling and water renewals, and aging of organisms. Additional intercalibrations are recommended specifically for the *Ceriodaphnia* reproduction test to assess sources of variability in both stormwater and laboratory dilution water.

TABLE OF CONTENTS

1 INTRODUCTION

Municipal stormwater monitoring programs in southern California are different than most other monitoring programs around the United States. The southern California monitoring programs differ because stormwater managers have invested in the Stormwater Monitoring Coalition (SMC), a consortium of all the primary regulated and regulatory stormwater agencies overseeing more than 5,000 stream miles (Table 1, www.SoCalSMC.org). Although the consortium consists of at least seven distinct local monitoring programs, the SMC has established a continuing goal to compile local monitoring data to make region-wide assessments.

Table 1. Members of the Southern California Stormwater Monitoring Coalition.

SMC Member Agency
Los Angeles County Flood Control District
County of Orange, OC Public Works
County of San Diego, Department of Public Works
Riverside County Flood Control and Water Conservation District
San Bernardino County Flood Control District
Ventura County Watershed Protection District
City of Long Beach Public Works Department
City of Los Angeles, Department of Public Works
California Regional Water Quality Control Board, Santa Ana Region
California Regional Water Quality Control Board, Los Angeles Region
California Regional Water Quality Control Board, San Diego Region
State Water Resources Control Board
California Department of Transportation
Southern California Coastal Water Research Project

In order to compile local monitoring programs into regional assessments, the SMC has expended considerable effort to design monitoring programs with similar goals and objectives, integrated sampling efforts, establishing standardized data protocols, and focused training opportunities (SMC BWG 2007). However, none of the SMC agencies have their own analytical laboratories. Inventories of analytical efforts among regional contract laboratories indicated differences in laboratory methods and this raised concerns amongst SMC member agencies about data comparability.

In 2003, the SMC launched their first laboratory intercalibration study to help ensure analytical comparability. The intercalibration focused on chemical measurements, established common reporting levels and target analytes, utilized iterative round robin exercises to minimize inter-laboratory variation, ultimately setting limits for precision and accuracy for SMC monitoring. The project culminated in a performance-based laboratory guidance manual (Gossett et al. 2004). The intercalibration was so successful, the SMC repeated the intercalibration two more times (Gossett and Schiff 2007, 2010). The comparability observed in the later intercalibrations rivaled the first intercalibration, indicating some residual memory in the system. This is especially good news because system memory would result in consistently high quality data during the intervening years. The success of the chemistry intercalibration exercises was primarily due to three factors: 1) communication and commitment among laboratory

personnel; 2) setting performance-based criteria for establishing standards of success; and 3) using locally derived reference materials including using a stormwater matrix.

Based on the success of the chemistry intercalibrations, the SMC has decided similar inter-laboratory variability evaluations and steps towards comparability should be taken for aquatic toxicity. All of the SMC member agencies conduct aquatic toxicity both for their individual regulatory-based monitoring requirements, as well as their collaborative, integrated regional monitoring. Cumulatively, the SMC spends nearly $1 million annually on laboratory toxicity testing, not including field sampling. Like chemistry, the need for comparability in toxicity testing among SMC member agencies remains a priority as managers evaluate the regional extent and magnitude of toxicity, compare toxicity between watersheds, or assess changes in toxicity over time as management actions are implemented.

1.1 OBJECTIVES AND GOALS OF THIS DOCUMENT

The objective of this toxicity testing guidance manual is to enhance comparability by presenting the performance-based guidelines established during the SMC toxicity laboratory intercalibration. This document makes recommendations for species selection based on a set of consistency criteria and presents already-established guidelines for methods. Also, the manual presents additional information on methods standardization to enhance comparability for SMC projects. Finally, the goal of the document is to quantify the level of comparability observed during the intercalibration, setting the current best-practices standards for minimum expectations of within (intra-) and between (inter-) laboratory variability for laboratories in the region testing actual runoff samples.

Although every laboratory involved in the toxicity intercalibration study was certified by the State of California Environmental Laboratory Accreditation Program (ELAP), inventories of each laboratory's protocols demonstrated that most are not using exactly the same procedures, and the federally- or state-approved methods allow for this flexibility. This study documents and quantifies the variability in toxicity testing, and provides some suggestions for reducing this inter-laboratory variability.

This study was guided by an Advisory Committee composed of toxicity laboratory managers throughout California (Table 2). The Advisory Committee included five contract laboratories (including all of the laboratories currently testing SMC runoff samples), two university laboratories, and three municipal laboratories. Cumulatively, this Advisory Committee exceeded 200 person-years of toxicity testing experience.

This guidance manual is a living document. It should be revisited each time an intercalibration exercise is conducted and can be expanded to include additional species, additional laboratories, or to refine the precision expectations as new information becomes available.

This document and laboratory intercalibration study is not a certification program and it does not circumvent regulatory requirements. The guidelines set by this document merely express the desired needs of stormwater agencies throughout the southern California region. Therefore, these stormwater agencies can use these guidelines in establishing specifications for work assignments or requesting proposals to conduct stormwater analyses. Alternatively, or in combination, stormwater regulatory agencies may use these specifications in the development of regulatory expectations for laboratory performance by monitoring agencies.

Table 2. Participating Laboratories and members of the Advisory Committee for the Stormwater Monitoring Coalition Toxicity Intercalibration.

Laboratory	**Advisory Committee Member**
Aquatic Bioassay & Consulting	Joe Freas
Aquatic Toxicity Laboratory (University California Davis)	Linda Deanovic
Aquatic Testing Laboratories	Joe LeMay
City of Los Angeles	Stan Asato
City of San Diego	Nick Haring
County Sanitation Districts of Los Angeles County	Christina Pottios
Marine Pollution Studies Laboratory at Granite Canyon (University California Davis)	Bryn Phillips
MBC Applied Environmental Sciences	Sonja Beck
Nautilus Environmental	Marilyn O'Neill
Pacific EcoRisk	Stephen Clark

2 GUIDANCE INFORMATION

This guidance document consists of five elements. First is guidance for test species. Second is guidance for testing methods. Third is guidance for recommended standardization. Fourth is guidance for precision expectations for toxicity testing by species and endpoint, including the laboratory evaluation criteria and results from the intercalibration study. Fifth is recommended guidance for future intercalibration studies.

2.1 SPECIES GUIDANCE

Eighteen different species/endpoint combinations are tested amongst the various SMC monitoring programs based upon requirements listed in their NPDES permits (Table 3). In preparation for the intercalibration study, the Advisory Committee identified six criteria for selecting species and endpoints for the intercalibration, which also provides guidance for ongoing monitoring:

- Commonly tested organism to provide spatial and temporal consistency
- Species sensitivity to provide changes in magnitude of effect
- Ease of testing to minimize cost
- Species availability to ensure test feasibility
- Representative of local species to provide environmental relevance
- Sufficient number of laboratories capable of conducting the test

Based on these criteria, the Advisory Committee prioritized four species, two freshwater and two marine, for intercalibration testing (Table 4).

2.2 METHOD GUIDANCE

Guidance for toxicity testing is provided by the US EPA and/or the State of California. Tables 5 to 8 present testing method guidance published in standardized manuals using approved protocols. The only exception is *Hyalella*, where the SMC guidance closely follows the guidance provided by the California Surface Water Ambient Monitoring Program.

Table 3. Inventory of stormwater toxicity testing by matrix, species and endpoints for SMC member agencies, 2013.

Agency	Freshwater		Marine	
	Species	**Endpoints**	**Species**	**Endpoints**
Ventura Co.	Inland silverside	Survival / Growth	Topsmelt	Survival / Growth
	Fathead Minnow	Survival / Growth	Kelp alga	Germination / tube elongation
	Ceriodaphnia	Survival / Reproduction	Purple sea urchin	Development
	Green alga	Biomass		
Los Angeles Co.	*Ceriodaphnia*	Survival / Reproduction	Purple sea urchin	Fertilization
City of Long Beach	Fathead Minnow	Survival / Growth	Topsmelt	Survival / Growth
	Ceriodaphnia	Survival / Reproduction	Kelp alga	Germination / tube elongation
	Green alga	Biomass	Purple sea urchin	Development
Orange Co.	Amphipod *Hyalella*	Survival / Growth	Purple sea urchin	Development
	Ceriodaphnia	Survival / Reproduction	Mysid	Survival / Growth
	Fathead Minnow	Survival	Purple sea urchin	Fertilization / Development
Riverside Co.	Fathead Minnow	Survival / Growth		
	Ceriodaphnia	Survival / Reproduction		
	Green alga	Biomass		
	Amphipod *Hyalella*	Survival / Growth		
San Diego Co.	Amphipod *Hyalella*	Survival / Growth	Purple sea urchin	Development
	Ceriodaphnia	Survival / Reproduction		
	Fathead Minnow	Survival		
City of Los Angeles	Fathead Minnow	Survival / Growth	Red abalone	Development
	Ceriodaphnia	Survival / Reproduction	Topsmelt	Survival / Growth
	Green alga	Biomass	Kelp alga	Germination / tube elongation

Table 4. Recommended species and endpoints for Stormwater Monitoring Coalition intercalibration toxicity testing

Common Name	Scientific Name	Test Duration	Endpoints
Ceriodaphnia	*Ceriodaphnia dubia*	6-8-d	Survival/reproduction
Amphipod	*Hyalella azteca*	96-h	Survival
Purple sea urchin	*Strongylocentrotus purpuratus*	72-h	Development
Mussel	*Mytilus galloprovincialis*	48-h	Development

Table 5. Recommended test conditions and test acceptability criteria for *Ceriodaphnia dubia*, survival and reproduction toxicity tests (EPA test method 1002.0, EPA 2002)

Test Type:	Static renewal (required)
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)
Light quality	Ambient laboratory illuminations (recommended)
Light intensity	10-20 µE/m2/s, or 50-100 ft-c
Photoperiod	16 h light, 8 h dark (recommended)
Test Chamber size	30 mL (recommended minimum)
Test solutions volume	15mL (recommended minimum)
Renewal of test solutions	Daily (required)
Age of test organisms	Less than 24 h; and all released within a 8-h period (required)
No. neonates per test chamber	Assigned using blocking by known parentage (required)
No. replicate test chambers per concentration	10 (required minimum)
No. neonates per test concentration	10 (required minimum)
Feeding regime	Feed 0.1 mL each of YCT and algal suspension per test chamber daily (recommended)
Cleaning	Use freshly cleaned glass beakers or new plastic cups daily (recommended)
Aeration	None (recommended)
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 Dilute Mineral Water
Test concentrations	Effluents: 5 and a control (required minimum) Receiving Water: 100% receiving water (or minimum of 5) and a control (recommended)
Dilution factor	Effluents: ≥0.5 (recommended) Receiving Waters: None or ≥ 0.5 (recommended)
Test duration	Until 60% or more of surviving control females having three broods (maximum test duration 8 days) (required)
Endpoints	Survival and reproduction (required)
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)
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 test, a minimum of three samples (e.g. collected on days one; three, and five) with a maximum holding time of 36h before first use (Required)
Sample volume required	1 L/day (recommended)

Table 6. Recommended test conditions and acceptability criteria for conducting *Hyalella azteca* 96-hour water only tests (adapted from SWAMP 2008 and EPA 2002)

Parameter	Conditions
Test type	Water-only test
Dilution series	Control and 5 test concentrations (0.5 dilution factor)
Temperature	23 ± 1°C
Light quality	Wide-spectrum fluorescent lights
Illuminance	About 50-100 fc
Photoperiod	16L:8D
Renewal of water	Renewal at 48 h
Age of Organisms	*H. azteca:* 7- to 14-d old (1- to 2-d range in age)
Test chamber	Plastic cups or glass beakers (covered with glass or plastic)
Volume of water	Minimum 100 ml per replicate
Number of organisms/chamber	5
Number of replicate chambers/treatment	6
Feeding	1 mL YCT (1800 mg/L stock) at 48 h, at least 2 hours prior to renewal
Substrate	1 square inch of nitex screen in each test chamber
Aeration	None; unless dissolved oxygen (DO) falls below 4.0 mg/L or the sample has high likelihood to have increased biological oxygen demand (BOD)
Dilution water	Laboratory water with a hardness between 80-200 mg/L CaCO3
Water quality	Temperature daily
	Conductivity, dissolved oxygen, and pH at least at the beginning of a test, at the 48hr renewal, and at test termination
	Alkalinity and hardness on undiluted sample and control/dilution water only
Test duration	96 +/- 2hr
Endpoints	Survival
Test Acceptability	90% control survival

Table 7. Recommended test conditions and acceptability criteria for conducting *Strongylocentrotus purpuratus* embryo development tests (EPA/600/R-95/136, EPA 1995)

Parameter	*Criteria*
Test Type	Static non-renewal
Salinity	34 ± 2 ‰
Temperature Range	15 ± 1.0 °C
Light Quality	Ambient laboratory illumination
Light Intensity	10-20 µE/m²/s
Photoperiod	16 hours light, 8 hours dark (ambient laboratory levels)
Test Chamber Size	30 mL
Test Solution Volume	10 mL
No. replicate chambers per concentration	4
Dilution Water	Uncontaminated 1-µm filtered natural seawater or hyper-saline brine prepared from natural seawater
Test Duration	72 ± 2 hr
Endpoints	Normal development, mortality can be included
Test Acceptability Criteria	>80% normal shell development in controls; must achieve a %MSD of <25%
Minimum Sample Volume	1 L per test

Table 8. Recommended test conditions and acceptability criteria for conducting *Mytilus galloprovincialis* embryo development tests (EPA/600/R-95/136, EPA 1995)

Parameter	*Criteria*
Test Type	Static non-renewal
Salinity	34 ± 2 ‰
Temperature Range	15 or 18 ± 1.0 °C
Light Quality	Ambient laboratory illumination
Light Intensity	10-20 µE/m²/s
Photoperiod	16 hours light, 8 hours dark (ambient laboratory levels)
Test Chamber Size	30 mL
Test Solution Volume	10 mL
No. replicate chambers per concentration	4 (plus 3 chemistry vials)
Dilution Water	Uncontaminated 1-µm filtered natural seawater or hyper-saline brine prepared from natural seawater
Test Duration	48 hr (or until complete development up to 54 hours)
Endpoints	Survival and normal shell development
Test Acceptability Criteria	Control survival must be ≥ 50% in control vials; >90% normal shell development in surviving controls; and must achieve a %MSD of <25%
Minimum Sample Volume	1 L per test

2.3 STANDARDIZATION GUIDANCE

While SMC guidance recommends using approved test methods (Tables 5-8), sometimes the approved methods have options for test procedures. In this section, we describe further recommendations to refine approved procedures for enhancing comparability when testing SMC samples. These recommended procedures are based on the majority consensus of the Advisory Committee.

2.3.1 HARDNESS OF DILUTION WATER

The Advisory Committee recommends the use of moderately hard laboratory water for controls and dilution water in stormwater toxicity tests. The hardness should be within a range of 80-100 mg/L $CaCO_3$. There are a number of options for creating moderately hard dilution water utilizing EPA Methods (EPA-821-R-02-013, EPA 2002) including adding reagents (salts, macro- and micro-nutrients) to deionized water, or dilute mineral water. The SMC laboratories utilized both methods for the intercalibration study which did not appear to affect the variability associated with testing. The Advisory Committee did not make a recommendation for what hardness should be used for culturing organisms. Culture water controls can be tested concurrently with SMC samples if hardness is a concern.

2.3.2 FEEDING FOR *CERIODAPHNIA* TESTS

The Advisory Committee recommends that *Ceriodaphnia* be fed daily during testing using a combination of yeast, Cerophyll®, trout chow (YCT), and *Selenastrum.* Currently, EPA Methods (EPA-821-R-02-013, EPA 2002) allow for options in recipes and timing.

2.3.3 SIEVE FOR *HYALELLA* AGE CLASS

The Advisory Committee recommends that *Hyalella* test organisms be 8-10 days old. This is a narrower range than recommended by the SWAMP guidance of 7-14 days (SWAMP 2008). The rationale for the narrower age range is due to sensitivity of *Hyalella* to various toxicants at differing age classes. To achieve this size range, the Advisory Committee recommends that *Hyalella* specimens be selected based on a nested sieve size of 600 and 500 µm (in that order), using the organisms with the selected diameter retained on the 500 µm sieve. As test organisms are added to test chambers, they should be visually examined to ensure they are generally within 1,600 to 1,800 µm in length.

2.3.4 SAMPLE HANDLING AND WATER RENEWALS

The Advisory Committee recommends that samples not be manipulated, such as filtering or allowed to settle, prior to subsampling for testing. Instead, samples should be vigorously shaken for 60 seconds prior to subsampling for testing, and again prior to pouring into test containers, to ensure sample representativeness. Similar sample handling steps should occur for daily water renewals.

3 PRECISION GUIDANCE

There have been several studies to assess the variability both within and among laboratories for toxicity testing (EPA 2001, Warren-Hicks et al. 2000, Burton et al. 1996). However, none have been explicitly for stormwater matrix, particularly in southern California. The focus of this intercalibration exercise was to quantify this variability, and hopefully minimize it (see Section 2.3), for the recommended SMC toxicity testing species.

3.1 APPROACH AND METHODS

The approach to this intercalibration required five steps:

1) Recruit qualified toxicity testing laboratories into the study, and use the combined expertise of the laboratory managers to help design and score the study (Section 3.1)
2) Create and distribute homogenized samples, blind to each of the testing laboratories for round 1 of intercalibration testing (Section 3.2.1)
3) Collectively review results with the testing laboratories and look to improve comparability, where necessary (Section 2.3)
4) Create and distribute homogenized samples, blind to each of the testing laboratories for round 2 of intercalibration testing (Section 3.2.2)
5) Document the range of variability and expectations for precision in this guidance manual (Sections 3.3 and 4.0)

The list of participating laboratories by testing method is provided in Table 9. This list of participating laboratories includes five contract laboratories (including all of the laboratories currently testing SMC runoff samples), two university laboratories, and three municipal laboratories. Laboratories are listed anonymously, but a key is available to SMC members.

The study was comprised of two intercalibration rounds. The first intercalibration round was comprised of four sample types: Lab dilution water (LDW), copper spiked into lab dilution water (CS), simulated runoff (SR), and duplicate (DUP) copper spiked into lab dilution water. For the freshwater test species, the LDW exposure was very hard freshwater created according to EPA (2002), using reconstituted deionized water. The LDW used for the CS was moderately hard freshwater created according to EPA (2002), using reconstituted deionized water. The nominal copper spiking concentration was 60 µg/L for *Ceriodaphnia* and 200 µg/L for *Hyalella*. No chemical confirmation was conducted for round 1. Lab dilution water for the marine species was 0.45 µm, charcoal filtered seawater collected from Santa Monica Bay. Nominal copper spiking concentration was 40 µg/L for both *Strongylocentrotus* and *Mytilus*. The SR samples for both freshwater and marine test species were collected by washing down approximately 400 m^2 parking lot with 200 L of activated carbon filtered tap water. No chemistry was conducted on the SR sample. For the marine species, each testing laboratory adjusted samples for salinity prior to testing using artificial sea salts.

With only one exception, all laboratories began test initiation on the same day. Each laboratory created a dilution series for each sample (100%, 50%, 25%, 12.5%, 6.25% sample, plus control). An effects

concentration at 25% (EC25) was estimated for each dilution series, when possible. The EC25 was calculated using CETIS© software (Tidepool Scientific, McKinleyville, CA).

Table 9. Species testing by participating laboratories in the SMC Toxicity Intercalibration.

	Round 1				Round 2[b]			
Laboratory	**CD[a]**	**HA[a]**	**SP[a]**	**MG[a]**	**CD**	**HA**	**SP**	**MG**
A	X	X	X	-	X	X		
B	X	X	-	-	X	X		
C	X	X	-	-	X	X		
D	-	-	-	-	-	-		
E	X	-	-	X	-	-		
F	X	X	X	X	X	X		
G	X	-	-	-	-	-		
H	X	X	X	X	-	-		
I	X	X	X	X	X	X		
J	-	X	X	X	X	X		

[a] CD=*Ceriodaphnia dubia*, HA=*Hyalella azteca*, SP=*Strongylocentrotus purpuratus*, MG=*Mytilus galloprovincialis*.
[b] No marine tests were conducted during round 2.

The second intercalibration round was also comprised of four sample types: LDW, CS, SR, and DUP using LDW. For the freshwater test species, the LDW was moderately hard freshwater created according to EPA Method EPA-821-R-02-013, using reconstituted deionized water. Nominal copper spiking concentration was 60 µg/L for *Ceriodaphnia* and 200 µg/L for *Hyalella*. Chemical confirmation was conducted for round 2; dissolved copper concentrations in samples collected from the master batch prior to sample split was 59.8 µg/L for *Ceriodaphnia* and 195 µg/L for *Hyalella*. The SR samples were collected by washing down approximately 400 m^2 parking lot with 200 L of activated carbon filtered tap water. Chemistry was conducted on the simulated runoff sample for total suspended solids (TSS) from each laboratory, specifically to assess the homogeneity of the SR sample among laboratories. TSS concentrations for *Hyalella* tests ranged from 25 to 83 mg/L with a coefficient of variation among laboratories of 38%. TSS concentrations for *Ceriodaphnia* tests ranged from 63 to 91 mg/L with a coefficient of variation of 14%. No marine species were tested in round 2.

Each sample was scored for comparability based on three factors including test acceptability, intra-laboratory precision, and inter-laboratory precision. See Appendix A for the complete scoring protocol developed by the Advisory Committee. Briefly, for each species, labs could receive a total of 12 points for test acceptability, 48 points for within laboratory precision, and 48 points for among laboratory precision. From a possible total of 108 points, four categorical assignments of comparability were assigned:

>90% of points = Very Highly comparable
>80% of points = Highly comparable
>70% of points = Moderately comparable
<70% of points = Low comparability

A cutoff between comparable and not comparable was at 70% (between moderate and low comparability).

Test acceptability was scored based on acceptable control survival and reference toxicant response. Intra-laboratory precision was based on the relative percent difference (RPD) between blind DUP samples analyzed within each laboratory. Inter-laboratory precision was based on the absolute difference between the laboratory's result and the grand mean result of all laboratories. The EC25 was used to estimate inter-laboratory precision, unless at least one laboratory could not achieve an EC25 (insufficient toxic dose response), in which case the percent effect in the 100% undiluted sample was used. Before calculating the grand mean, laboratory results were screened for outliers using the Grubbs test.

Most inter- and intra-laboratories comparisons made in this section of the report are examining ranges of effect levels in 100% sample, since this is the primary comparison that will be used by SMC member agencies. However, many other comparisons have been used by others in past toxicity intercalibrations including coefficient of variation (CV) (EPA 2001), relative percent difference (RPD), *h* statistic and *k* statistic (Burton et al. 1996). The test of significant toxicity (TST) was specifically not used for comparing test outcomes.

3.2 INTRA- AND INTER-LABORATORY PRECISION

3.2.1 ROUND 1 INTERCALIBRATION

All of the responses for each test endpoint to 100% sample in Round 1 are presented in Figure 1. Examining each endpoint, some samples exhibited more comparable responses than others. For example, *Mytilus* had very similar responses to 100% sample for all four samples. Excluding one outlier for the stormwater sample, laboratories differed by no more than 5% effect for any sample. In contrast, *Ceriodaphnia* survival varied the most among 100% samples in Round 1. Laboratory responses ranged from 40% effect for LDW to 100% effect for the CS sample. The range of laboratory responses for the CS Dup was nearly as large as the CS (80%). Interestingly, the average response among laboratories between the CS and CS Dup was similar (61% vs. 63% effect), and the average RPD within laboratory was relatively low (6%), indicating laboratories can reproduce their own data even when inter-laboratory variability is large.

There were two noteworthy results from Round 1 (Figure 1). The first noteworthy result was the relative precision for the SR sample versus all other samples. Regardless of test endpoint, the SR sample variation among laboratories was less than, or similar to, the LDW, CS, or CS DUP samples. This may be due, in part, to the sensitivity of the various endpoints. SR samples were either highly toxic (100% effect in *Strongylocentrotus*, *Mytilus*, *Ceriodaphnia* reproduction) or non-toxic (close to 0% effect *Ceriodaphnia* survival, *Hyalella*). The second noteworthy result was the variation among laboratories for the LDW. This sample of dilution water, which was prepared using standard methods, elicited toxic responses (up to 60% effect) from multiple laboratories.

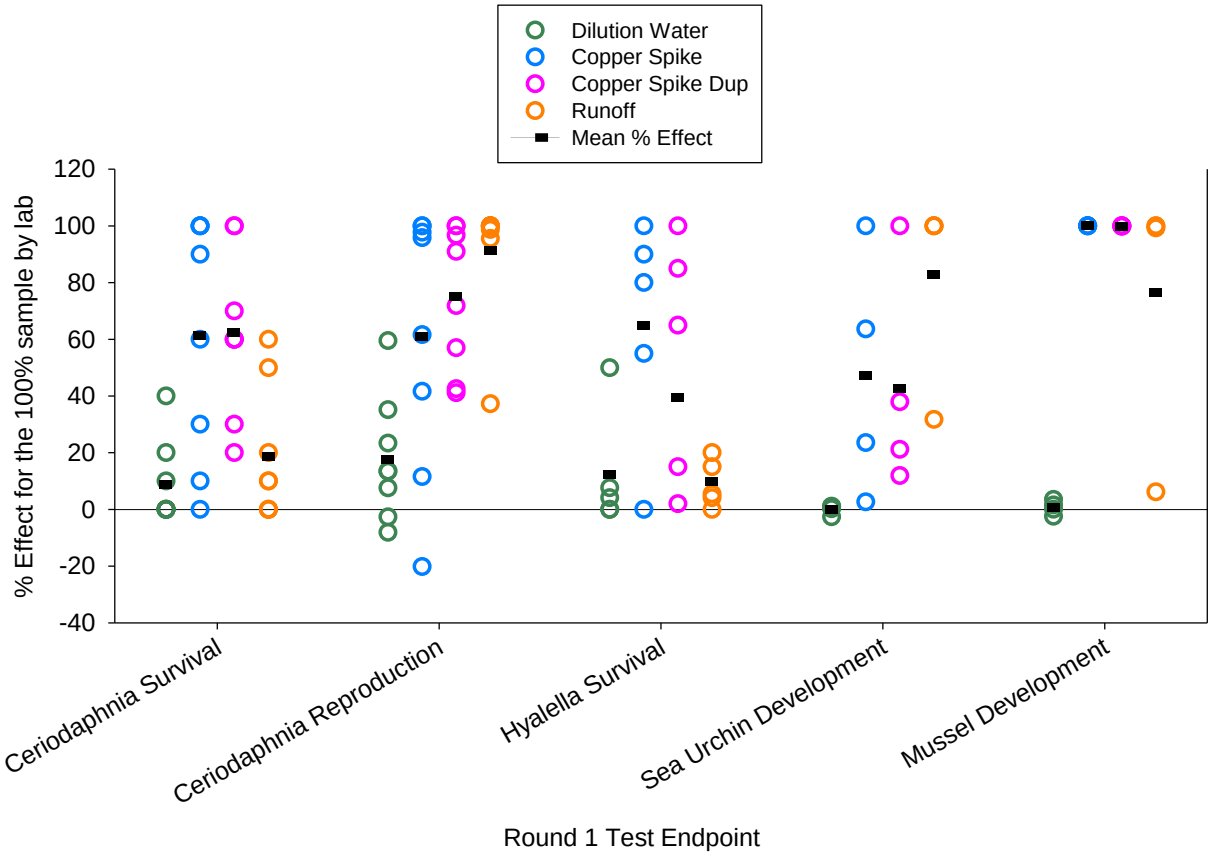

Figure 1. Toxicity test response (% effect) of the various endpoints to full strength (no dilution or 100%) samples during Round 1 of the SMC intercalibration study. Each symbol represents the result from a single laboratory (see Table 9).

3.2.1 Round 2 Intercalibration

All of the responses for each test endpoint to 100% sample in Round 2 are presented in Figure 2. In general, the comparability amongst laboratories for both the *Ceriodaphnia* survival and *Hyalella* endpoints improved between Round 1 and Round 2. For *Hyalella*, the LDW and LDW DUP were not toxic, and the range of toxicity observed for the 100% SR sample ranged from 10% to 40% effect. For *Ceriodaphnia* survival, the LDW and LDW DUP were almost universally <20% effect, and the range of toxicity observed for the 100% SW sample was <30% effect.

The interlaboratory precision improvements observed for *Hyalella* were particularly noteworthy. The Advisory Committee attributed these improvements to the standardization among laboratory protocols (Table 6) and the minimization of age range (Section 2.3.3). Both of these standardization techniques were implemented between round 1 and round 2.

In contrast to *Hyalella*, interlaboratory precision for *Ceriodaphnia* reproduction did not demonstrably improve between round 1 and round 2. The toxicity observed in round 2 LDW and LDW DUP ranged >70% effect and approximately half the laboratories observed toxicity >20% effect. For *Ceriodaphnia* reproduction, the range of toxicity observed for the 100% SR sample in round 2 was approximately 50%

effect, greater than the range observed in round 1. For this endpoint, the standardization steps in Section 3.2.3 did not appear to improve consistency and comparability among laboratories.

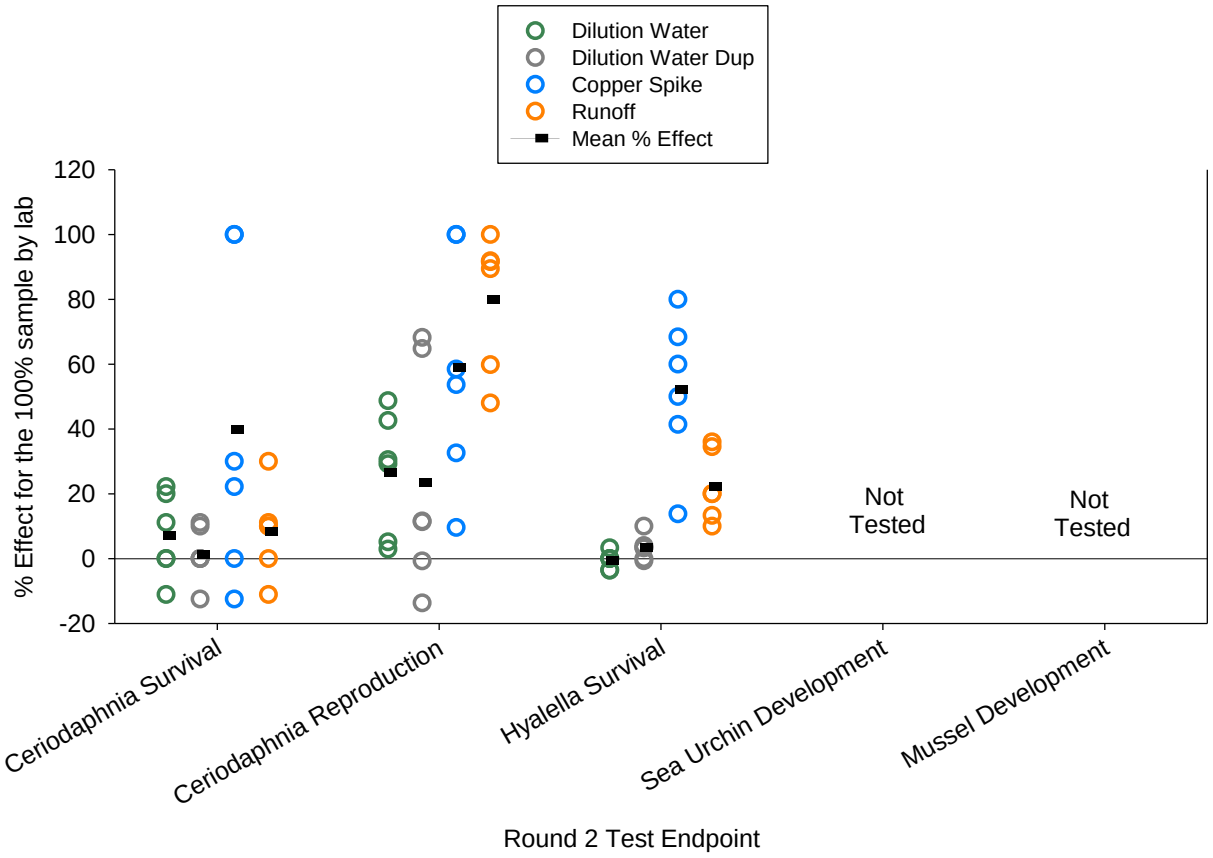

Figure 2. Toxicity test response (% effect) of the various endpoints to full strength (no dilution or 100%) samples during round 2 of the SMC intercalibration study. Each symbol represents the result from a single laboratory (see Table 9).

3.3 SUMMARY OF COMPARABILITY

In general, SMC member agencies can expect a range of results from various labs testing the same runoff sample, which varies by test organism and endpoint. Based on the results from this study, the range of effect concentrations in 100% simulated runoff samples were *Ceriodaphnia* survival (ca. ± 10% effect), *Hyalella* survival (ca. ± 12% effect), *Ceriodaphnia* reproduction (ca. ± 22% effect), *Strongylocentrotus* development (ca. ± 40% effect, 2% effect without outliers), and *Mytilus* embryo development (ca. ± 45% effect).

Based on the scoring system developed for this study, the participating laboratories were comparable for most of the test endpoints (Table 10). Virtually all laboratories were able to meet test acceptability requirements, including internal positive and negative controls. Most laboratories tended to produce internally consistent results when given blind duplicate samples. Finally, most laboratories produced data consistent with non-toxic samples when exposed to laboratory dilution water.

Table 10. Summary of laboratory comparability scoring for *Ceriodaphnia dubia* (6-8 day) survival and reproduction, *Hyalella* survival, *Strongylocentrotus* embryo development, or *Mytilus* embryo development tests.

Lab	*Ceriodaphnia* Survival		*Ceriodaphnia* Reproduction		*Hyalella* Survival		*Strongylo-centrotus* Development	*Mytilus* Embryo Development
	Round 1	*Round 2*	*Round 1*	*Round 2*	*Round 1*	*Round 2*	*Round 1* [a]	*Round 1* [a]
A	Moderate	High	Very High	Low	Low	High	Moderate	- [b]
B	Very High	High	Moderate	High	Low	High	-	-
C	Low	High	Low	High	Low	Very High	-	-
E	Moderate	-	Moderate	-	-		-	Very High
F	Moderate	High	Moderate	Low	Low	Very High	Moderate	Low
G	High	-	High	-	-	-	-	-
H	Low	-	Low	-	-	-	-	-
I	High	Moderate	High	Low	Moderate	Very High	High	Very High
J	Low	High	Low	Low	High	Very High	Moderate	Moderate

[a] Only tested in Round 1
[b] – indicates sample not tested

The primary exception to this trend was for the *Ceriodaphnia* reproduction test. Although inter-laboratory variability was consistently lowest when testing simulated runoff samples, inter-laboratory variability increased for both lab dilution water and copper spiked lab dilution water resulting in a wide range of comparability scoring for this endpoint. In round 1 and round 2, half of the laboratories were deemed comparable, but the laboratories deemed comparable differed between round 1 and round 2; low comparability labs in round 1 improved comparability scores in round 2 and vice-versa. This lack of consistency may indicate a variable test method, the need for an improved study design, or both.

The amount of testing variability observed during this intercalibration for *Ceriodaphnia* is not uncharacteristic of the variability observed by others examining wastewater effluents, reference toxicants, or ambient media. Moore et al. (2000) used split samples of lab dilution water among 16 laboratories for *Ceriodaphnia* reproduction tests and observed a mean response of 16% effect and a standard deviation of 28% effect. This is somewhat comparable to the variability we observed for laboratory dilution water split samples, which had a response that ranged from 16 to 27% effect, and a standard deviation of 19 to 27% effect. Diamond, *et al*. (2008) used six laboratories to test split samples of lab dilution water for *Ceriodaphnia* reproduction, and all six had IC25s within 35% effect.

EPA (2001) conducted a thorough toxicity intercalibration for the *Ceriodaphnia* reproduction endpoint as part of its whole effluent testing (WET) program. The EPA intercalibration observed a mean response of 3% effect amongst 27 labs that tested laboratory dilution water; 26 of the 27 laboratories estimated an IC25 of >100. However, laboratory dilution water samples were prepared differently for the EPA intercalibration compared to the current study. Sealed ampules of laboratory dilution water were delivered to each of the testing laboratories, who were directed to dilute the ampule by 10:1 (similar to an EPA DMR Quality Assurance sample). Thus, maximum split sample concentration was 10% (EPA 2001).

In addition to lab dilution water, EPA (2001) also tested split samples of positive controls (spiked copper in lab dilution water) and wastewater effluents among the 27 laboratories analyzing *Ceriodaphnia* reproduction. Cumulatively, the coefficient of variation (CV) in IC25s across all 27 laboratories was

estimated at 35%. When the *Ceriodaphnia* reproduction results from the SMC are calculated in a similar fashion to EPA (2001), the CV was somewhat comparable at 39.5%. Differences in overall CVs could be a function of increased variability in test results, fewer testing laboratories (N=27 vs 6), or both.

There could be a number of reasons for the variability observed in *Ceriodaphnia* survival and reproduction tests in this SMC intercalibration. Virtually all of the participating labs were able to initiate tests and meet test acceptability criteria in the present study, but DeGraeve et al. (1992) had 44% of their intercalibration tests fail to initiate due to unsuccessful cultures or unacceptable neonate production during testing. EPA (2001) had 18% of their tests fail test acceptability criteria for similar reasons. Cooney et al. (1991) identified water renewal and feeding regime as primary factors influencing results in reference toxicity tests. LaRocca et al. (1994) and Belanger et al. (1989) also identified feeding regime as an important variable influencing *Ceriodaphnia* test results. Although the present study did not find a clear relationship between feeding and water renewals and test variability, both are part of the standardized guidance recommended in this manual (which is consistent with the standard protocol from EPA, section 2.2).

The evidence in the literature is mixed regarding the effect differences in hardness between culture water and laboratory dilution water might have on samples being tested. Belanger et al. (1989) did not detect a measureable difference in the toxicity of copper to *Ceriodaphnia* when tested in moderately hard dilution water, regardless if the organisms were cultured in hard or moderately hard water. In contrast, Naddy et al. (2003) did find an effect of copper to *Ceriodaphnia* when cultured in hard versus moderately hard water, but tested in moderately hard water. Although the present study did not find a clear relationship between hardness and test variability, it is part of the standardized guidance recommended in this manual (which is consistent with the standard protocol from EPA, section 2.2).

4 RECOMMENDATIONS FOR IMPROVING PRECISION

Understanding laboratory variability for toxicity testing proved to be more difficult than for chemistry. This may be due in part to the variability inherent in the test (i.e., live organisms), or it may be the variability inherent to the laboratories. Like chemistry, however, much was learned from this intercalibration study that provides recommendations for future efforts. These recommendations include improved study designs, testing frequency, and onboarding new laboratories.

4.1 FUTURE INTERCALIBRATION STUDY DESIGNS

The Advisory Committee recommends two improvements to the enhance future study designs to discern differences within or between laboratories conducting stormwater toxicity testing. The first study design enhancement is to increase the sample size for runoff samples, which was the primary goal of the study. In this study, only one runoff sample was distributed per round of testing. A more accurate assessment of intra- and inter-laboratory comparability would be achieved if more runoff samples were tested.

To accommodate the increased sample size, the Advisory Committee recommends dropping dilution series. The rationale for discontinuing dilution series is that undiluted sample concentrations were most frequently compared during this study and the SMC's emphasis of testing runoff samples without dilution series during routine monitoring.

The additional runoff samples recommended for future intercalibrations should come from a variety of watersheds, including presumed or historically toxic and non-toxic sites. Additional duplicate samples within the range of expected effects should also be included.

The second opportunity to enhance future intercalibration study designs is to focus on the *Ceriodaphnia* reproduction test, specifically for laboratory dilution water (LDW). During the present intercalibration, multiple laboratories observed toxicity in LDW prepared by another lab. The reason for the observed toxicity in a theoretically non-toxic sample is still unknown. Future study designs should confirm this anomalous result, conduct the experimental manipulations to identify the source of this inter-laboratory variability, then re-test the intercalibration to ensure any standardization actually improves inter-laboratory precision. Sample concentrations should also be chemically confirmed. The management impacts of variability with LDW is not unique to the SMC, and the new study design should incorporate all regulated discharges and regulatory agencies that rely on toxicity testing for management responses such as permit compliance, toxicity identification evaluations, or total maximum daily loads.

4.2 INTERCALIBRATION FREQUENCY

The Advisory Committee recommends that the intercalibration frequency mirror the chemistry intercalibration frequency of at least every three years. This will ensure confidence in results between intercalibrations including staff turnover and protocol modifications.

4.3 ONBOARDING NEW LABORATORIES

Since SMC member agencies will want to use laboratories that participated in the intercalibration, the Advisory Committee recognized that onboarding new laboratories that did not participate in these intercalibrations may be problematic. If a new laboratory wishes to test SMC samples, the Advisory Committee recommends that this laboratory follow the test procedures outlined in section 4.1; four samples delivered blind. The samples should be concurrently tested by one of the laboratories that passed the comparability testing in this intercalibration exercise as a reference lab. Successful onboarding for each species would occur if the new lab was within the range of laboratory variability indicated in Figures 1 and 2. Then, the new laboratory would be mandated to participate in the next SMC intercalibration study.

5 REFERENCES

Belanger, S.E., Farris, J.L. and Cherry, D.S., 1989. Effects of diet, water hardness, and population source on acute and chronic copper toxicity to *Ceriodaphnia dubia*. *Archives of Environmental Contamination and Toxicology*, 18(4), pp.601-611.

Burton, G.A., Norberg-King, T.J., Benoit, D.A., Ankley, G.T., Winger, P.V., Kubitz, J., Lazorchak, J.M., Smith, M.E., Call, D.J., Ingersoll, C.G. and Greer, E., 1996. Interlaboratory study of precision: *Hyalella azteca* and *Chironomus tentans* freshwater sediment toxicity assays. Environmental toxicology and chemistry, 15(8), pp.1335-1343

Cooney, J.D., Lenoble, B.J., Pollock, T.L., Smith, G.J., Degraeve, G.M. and Moore, E.L., 1992. Effects of environmental and experimental design factors on culturing and toxicity testing of *Ceriodaphnia dubia. Environmental toxicology and chemistry*, 11(6), pp.839-850.

DeGraeve, G.M., Cooney, J.D., Marsh, B.H., Pollock, T.L. and Reichenbach, N.G., 1992. Variability in the performance of the 7-d *Ceriodaphnia dubia* survival and reproduction test: An intra-and interlaboratory study. *Environmental toxicology and chemistry*, 11(6), pp.851-866.

Diamond, J, P Stribling, M Bowersox, 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

EPA. 1995. Short term methods for estimating the chronic toxicity of effluents and receiving water to west coast marine and estuarine organisms. United States Environmental Protection Agency, Office of Research and Development. Washington, DC. EPA/600/R-95/136

EPA. 2001. Final report: Interlaboratory variability study of 12 short-term chronic and acute whole effluent toxicity test methods. United States Environmental Protection Agency, Office of Water. Washington, DC. EPA-821-B-1-04

EPA. 2002. Short term methods for estimating the chronic toxicity of effluents and receiving water to freshwater organisms, Fourth Edition. United States Environmental Protection Agency, Office of Research and Development. Washington, DC. EPA-821-R-02-013

LaRocca, C.A., Francisco, D.E. and DiGiano, F.A., 1994. Effects of diet on survival, reproduction, and sensitivity of *Ceriodaphnia dubia*. *Water environment research*, 66(7), pp.905-911.

Moore, T.F., Canton, S.P. and Grimes, M., 2000. Investigating the incidence of type I errors for chronic whole effluent toxicity testing using *Ceriodaphnia dubia*. *Environmental toxicology and chemistry*, 19(1), pp.118-122.

Naddy, R.B., Stern, G.R. and Gensemer, R.W., 2003. Effect of culture water hardness on the sensitivity of *Ceriodaphnia dubia* to copper toxicity. *Environmental toxicology and chemistry*, 22(6), pp.1269-1271.

Stormwater Monitoring Coalition Bioassessment Working Group. 2007. Regional monitoring of Southern California's coastal watersheds. Technical Report 539. Southern California Coastal Water Research Project. Costa Mesa, CA.

Gossett R, K Schiff. 2010. Stormwater Monitoring Coalition Laboratory Guidance Document (Third Edition). Technical Report 615. Southern California Coastal Water Research Project. Costa Mesa, CA.

Gosset R, KC Schiff. 2007. Stormwater Monitoring Coalition laboratory guidance document (Second Edition). Technical Report 521. Southern California Coastal Water Research Project. Costa Mesa, CA.

Gossett R, D Renfrew, K Schiff. 2004. Stormwater monitoring coalition laboratory guidance document. Technical Report 420. Southern California Coastal Water Research Project. Westminster, CA

SWAMP. 2008. Surface Water Ambient Monitoring Program Plan. State Water Resources Control Board, Sacramento, CA http://www.waterboards.ca.gov/water_issues/programs/swamp/

Warren-Hicks, W.J., Parkhurst, B.R., Moore, D.R., Teed, R.S., Baird, R.B., Berger, R., Denton, D.L. and Pletl, J.J., 2000. Assessment of whole effluent toxicity test variability: partitioning sources of variability. *Environmental toxicology and chemistry*, 19(1), pp.94-104

APPENDICES

SCORING SOP

ROUND 2:
STORMWATER TOXICITY TESTING LABORATORY INTERCALIBRATION

DATA ANALYSIS
STANDARD OPERATING PROCEDURES

December 4, 2015

CONTENTS

SCOPE

This study has been commissioned by the Stormwater Monitoring Coalition (SMC) to quantify stormwater-sample testing comparability among laboratories for toxicity methods. This intercalibration study will assess variability among laboratories, identify potential quality improvements, and improve comparability and consistency in toxicity measurements. The main objectives of this study are to: 1) conduct two round robin exercises to characterize and ultimately minimize inter-laboratory variability for testing marine and freshwater species, and 2) develop a manual to provide guidelines for storm water toxicity testing precision and sensitivity. The purpose of this document is to detail the procedures for analyzing the data produced during the intercalibration exercises for the second round of exposures. A separate document that outlines sampling and testing logistics will be concurrently submitted to participating laboratories.

APPROACH

A Laboratory Working Group composed of expert laboratory managers developed comparability evaluation criteria for data generated by participants during Testing Round 1. Two freshwater species will be tested during Round 2. The species and endpoints selected for testing are the chronic *Ceriodaphnia dubia* (cladoceran) survival and reproduction and the 96-h *Hyalella azteca* (amphipod) acute survival tests. The testing and sample collection approaches for the 2nd round have been described in the Logistics SOP distributed on November 13, 2015.

INTER-LABORATORY COMPARABILITY

Successful completion of this exercise by a laboratory will be evaluated based on two criteria:

I. Attainment of test acceptability criteria
II. Comparability among laboratories

I. Attainment of Test Acceptability Criteria

For a test to be considered valid for the intercalibration, the following acceptability criteria must be met:

A. All controls must meet protocol specific minimum test acceptability criteria (Table 1). Six points per species will be assigned to each laboratory for meeting the species specific acceptability criteria (Table 2).
B. A valid concurrent reference toxicant test must be run with each batch of organisms (criteria in Table 1 must also be met). The reference toxicant test will provide among other data, information to determine that test animals were exhibiting normal sensitivity to the provided test samples. These steps will be followed to conduct this analysis:
   a. The reference toxicant test data will be compared to historic data previously provided by each laboratory for the most recent test (up to 20 tests).
   b. The reference toxicant test LC_{50} or EC_{50} should fall within two standard deviations of the laboratory historical mean, to indicate normal sensitivity.
   c. Up to six points per species will be awarded for the successful completion of a reference toxicant (Table 2).

If test acceptability criteria for the samples tested is not met by a laboratory, their data will not be used for any of the subsequent analysis. However, a laboratory will be given the opportunity to retest if acceptability criteria is not met. A laboratory planning to retest must contact SCCWRP within 24h of knowing that a test failed the acceptability criteria. Then, SCCWRP will identify a referee laboratory that achieved an acceptable test and help to make the appropriate arrangements for retesting.

Subsequently, the failing laboratory and the referee laboratory will work with SCCWRP to determine acceptable conditions under which the retest will occur. If a laboratory's reference

toxicant test fails to meet acceptability in Table 1, and/or is not within two standard deviations from their historical mean, this will not disqualify or affect the sample data analysis.

Table 1. Test acceptability criteria.

Parameter	Test Method	
Test type:	**Chronic *Ceriodaphnia***	**Acute *Hyalella***
Endpoints:	Survival and Reproduction	Survival
Test Acceptability Criteria:	Mean control survival must be ≥ 80%, ≥60 % of surviving control females must produce 3 or more broods; mean number of offspring per surviving females must be ≥ 15 neonates;	Mean control survival must be ≥ 90%.

PMSD = Percent minimum significant difference

Table 2. Scoring categories for the test acceptability criteria and reference toxicant data.

Acceptability Criteria	Points Assigned
Meets protocol specific minimum test acceptability criteria (1.5 points per sample)	6
Does not meet protocol specific minimum test acceptability criteria [a]	0
Reference toxicant test LC_{50} or EC_{50} falls within 2 standard deviations of laboratory historical mean and meets test acceptability criteria	6
Reference toxicant test LC_{50} or EC_{50} does not fall within 2 standard deviations of laboratory historical mean and does not meet test acceptability criteria	0

a = Data will not be used on subsequent analyses

II. Comparability Among And Within Laboratories

Inter-laboratory comparability will be based on two scoring categories:

1) **Comparison of effects for a given sample**. This analysis will help us to determine inter-laboratory variability and precision. Precision will be the degree of mutual agreement among individual measurements (Taylor, 1987).
2) **Relative percent difference (RPD) between a sample and its duplicate**. Laboratory duplicates are subsamples of the original sample that will be prepared and analyzed as a separate sample. The laboratory duplicate analysis will also provide information on the precision of the analysis and on the intra-laboratory variability within samples.

~~The following analysis will also be conducted to investigate comparability among laboratories, but will not be scored:~~

~~**3) Evaluation of the TST and determination of magnitude of toxicity**. This step will provide further inter-laboratory variability information~~

1) Comparison of Effect Levels for a Given Sample

Data for each species tested with the four samples will be evaluated separately. A grand mean will be calculated for the 25 effect level (EC_{25}) generated by each laboratory. Laboratory comparability scores will be calculated based on the percent difference of the EC_{25} from the grand mean and scored accordingly (Table 3).

- SCCWRP will distribute data templates to participants. Participating laboratories are expected to provide percent effect, EC25, and TST data generated with the CETIS software (e.g., multiple comparison test and the linear interpolation test). In addition, the labs must submit their original CETIS results, a copy of their bench sheets and raw data (same templates as used in Round 1).

- It is possible that for a given sample, a laboratory will not be able to determine the EC_{25} because the sample is either non-toxic or highly toxic. If one or more laboratories cannot calculate an EC_{25}, the scoring for all laboratories will be conducted using the percent effect in the undiluted sample compared to control. If a sample is highly toxic, we will use the percent effect in the 6.25% dilution compared to control.

- We will score *C. dubia* endpoints separately.

To calculate the grand mean the following steps will take place:

a) Pool data from all participating laboratories, treating each sample type and species separately.
b) Remove outlier laboratory's data for each sample (do not include in grand mean calculation; see description below).
c) Calculate a grand mean using the individual EC_{25} values or percent effect between the sample dilution and the control.
d) Calculate for each laboratory the percentage point difference from the grand mean.

Percentage point difference = Absolute value (laboratory EC_{25} - grand mean EC_{25})

e) Assign points to each laboratory based on the percentage point difference between their individual result and the grand mean (Table 3).

A Grubb's test will be used to determine outlier laboratory results (Grubbs, 1969). If a laboratory's sample is identified as an outlier, the outlier data will be removed before the grand mean will be calculated. All point estimates will be calculated using Linear Interpolation and following the EPA decision three.

Table 3. Scoring categories for percent difference of the EC_{25} from the grand mean. Points assigned per sample type and per species.

Difference from Grand Mean*	Points Assigned
0-7.5%	12
>7.5-15%	9
>15-22.5%	6
>22.5-30%	3
>30%	0

* = The percentage point difference value for each laboratory and the grand mean.

2) Relative Percent Difference (RPD) between a Sample and its Duplicate

A single blind duplicate for one of the three samples being tested will be distributed to all participating laboratories. For this comparison we will calculate the relative percent difference (RPD) between the sample and its duplicate for each species and dilution tested with the formula:

$$RPD = \frac{Abs\ (Sample\ EC_{25} - Duplicate\ EC_{25}) \times 100}{Average\ of\ Sample\ and\ Duplicate\ EC_{25}}$$

Where, Abs = Absolute Value

The following procedures will take place during this analysis:

a. Calculate RPD for each laboratory for the $EC_{25,}$ by species
b. Compare each laboratory RPD to the thresholds in Table 4
c. Assign points to each laboratory

Table 4. Scoring categories for the RPD analysis.

RPD Results	Points Assigned
0-10%	48
>10-20%	36
>20-30%	24
>30-40%	12
>40%	0

~~**3) Evaluation of the TST and determination of magnitude of toxicity**~~

~~To determine whether a sample is toxic or not we will use the Test of Significant Toxicity (TST). This category will not be scored and will not be used in the final integration of comparison factors to evaluate laboratory comparability. Each dilution will be treated as a separate sample for the purpose of TST analysis. The use of multiple dilutions will ensure that toxic and non-toxic samples/dilutions are available for this analysis. In addition, we will evaluate the effectiveness of the TST in identifying toxicity compared to traditional statistical methods. The following steps will take place for this analysis:~~

~~a) Run the TST by sample type, dilution, and species~~

~~b) Determine if sample is toxic or not using TST results~~

~~c) Compare TST results among laboratories for each dilution in a given sample~~

~~If a sample fails the TST analysis in the 100% sample concentration, the sample will be categorized according to Table 5, but based on percent effect from control in the undiluted sample.~~

~~d) Compare the category results between the individual laboratory results~~

Table 5. Toxicity thresholds for toxic samples.

Non-toxic	Effect is < 20% relative to the control (acute test) or < 25% relative to the control (chronic test)
Moderate Toxicity	> 20 to 40% effect (acute test) or > 25 to 50% effect (chronic test)
High Toxicity	> 40% effect (acute test) or > 50% effect (chronic test)

INTEGRATION COMPARISON FACTORS

Laboratories will be scored separately for each test type and comparability will be assessed for each species tested. The summed points will be compared to a category to determine comparability. For example, each laboratory will be categorized as very highly comparable, highly comparable, moderately comparable and low comparability (Table 6). The low comparability category is considered to be unacceptable. A process for addressing laboratories in the low category will be determined later, if needed.

Table 6. Potential comparability categories among laboratories. Maximum points possible represents results for a single species.

Description	% of Maximum Possible Score	No. of Maximum Points Possible
Very high comparability	≥ 90	≥ 97
High comparability	≥ 80	≥ 86
Moderate comparability	≥ 70	≥ 75
Low comparability	<70	<75

REFERENCES

Frank E. Grubbs. 1969. Procedures for Detecting Outlying Observations in Samples. Technometrics. Vol. 1, No 11.

Taylor, J.K. 1987. Quality Assurance of Chemical Measurements. Lewis Publishers, Inc., Chelsea, MI. 328 pp.

USEPA. 2000. Understanding and accounting for method variability in whole effluent toxicity applications under the National Pollutant Discharge Elimination System. EPA 833-R-00-003.

USEPA. 2010. National pollutant discharge elimination system test of significant toxicity implementation document. Office of Water. EPA 833-R-10-003.

LABORATORY RESULTS

ROUND 1 TABLES

Ceriodaphnia Reproduction				
Lab	Sample	Dilution	% Control	SD
A	1	100	76.68	10.1434
		50	106.07	10.0089
		25	107.029	11.6833
		12.5	93.2907	6.10646
		6.25	103.195	6.03784
	2	100	38.3387	8.67948
		50	82.0021	6.32456
		25	91.5868	6
		12.5	86.262	4.12311
		6.25	95.8466	3.65148
	3	100	0	0
		50	11.4087	2.74368
		25	38.3929	6.38053
		12.5	81.25	5.9963
		6.25	91.5179	3.10018
	4	100	28.115	9.12627
		50	88.8179	5.71159
		25	86.901	4.41714
		12.5	91.6933	6.32543
		6.25	96.1661	6.55659
B	1	100	86.5385	9.10433
		50	105.288	4.99889
		25	116.346	4.02216
		12.5	107.692	6.7363
		6.25	117.788	2.59272
	2	100	65.8654	6.41266
		50	87.9808	4.90011
		25	115.865	3.38132
		12.5	113.942	3.093
		6.25	120.192	2.82843
	3	100	0	0
		50	27.3092	2.4404
		25	74.2972	6.73713
		12.5	85.1406	3.32666
		6.25	100.803	3.72529
	4	100	42.9719	8.71844
		50	82.7309	3.1693

		25	85.1406	3.08401
		12.5	78.7149	3.27278
		6.25	77.1084	3.45768
C	1	100	64.881	5.78216
		50	70.8333	8.06157
		25	103.571	5.87272
		12.5	75.5952	4.71522
		6.25	92.2619	8.35663
	2	100	0	0
		50	0	0
		25	23.3333	4.60072
		12.5	82	6.01941
		6.25	80.6667	7.21803
	4	100	0	0
		50	0	0
		25	37.2093	4.94862
		12.5	68.6047	4.93964
		6.25	81.9767	6.50555
E	1	100	86.5604	3.01846
		50	86.3326	5.02107
		25	92.0273	2.91357
		12.5	88.8383	7.73161
		6.25	95.4442	4.62961
	2	100	1.4742	1.89737
		50	81.5725	14.413
		25	103.931	3.26769
		12.5	100.246	1.8738
		6.25	102.948	3.95671
	3	100	4.43459	3.01846
		50	68.9579	7.65143
		25	80.9313	7.1686
		12.5	95.122	4.79467
		6.25	100.443	4.54728
	4	100	8.99358	6.37356
		50	98.9293	4.31535
		25	90.7923	9.05784
		12.5	102.57	4.06749
		6.25	98.2869	6.43687
F	1	100	108.122	4.78539
		50	105.584	4.28952
		25	123.35	4.94526

		12.5	122.843	4.26354
		6.25	112.183	6.26188
	2	100	0	0
		50	0	0
		25	0	0
		12.5	103	6.39792
		6.25	114	5.41192
	3	100	62.7615	2.66667
		50	63.1799	3.72529
		25	73.2218	3.83695
		12.5	107.113	5.08156
		6.25	98.7448	4.06065
	4	100	0	0
		50	0	0
		25	0	0
		12.5	91.3978	5.61743
		6.25	106.989	5.64604
G	1	100	102.682	3.32666
		50	114.176	1.81353
		25	111.494	1.59513
		12.5	104.981	1.57762
		6.25	103.831	2.72641
	2	100	2.23881	1.89737
		50	85.4478	7.7093
		25	91.791	7.3967
		12.5	101.493	1.68655
		6.25	101.119	2.46982
	3	100	0	0
		50	15.5235	2.62679
		25	84.1155	5.18652
		12.5	110.83	4.47338
		6.25	106.859	1.83787
	4	100	3.2967	2.84605
		50	77.2894	5.83952
		25	86.0806	5.46199
		12.5	100.366	3.09839
		6.25	102.93	2.18327
H	1	100	92.4444	3.73571
		50	90.6667	8.08565
		25	107.556	7.23878
		12.5	101.778	3.38132

		6.25	106.222	5.17365
	2	100	88.4444	8.59522
		50	119.111	3.04777
		25	121.778	3.68782
		12.5	101.778	3.98469
		6.25	97.3333	1.79196
	3	100	1.2987	0.94868
		50	21.645	3.77124
		25	57.5758	4.54728
		12.5	107.792	6.87103
		6.25	113.42	4.70933
	4	100	57.4219	8.24688
		50	94.5313	9.11409
		25	119.531	6.65332
		12.5	84.375	6.7363
		6.25	87.5	8.66923
I	1	100	40.4762	10.6687
		50	71.4286	14.7874
		25	96.7262	2.83823
		12.5	92.8571	7.0206
		6.25	89.881	11.1833
	2	100	58.3333	14.8788
		50	105.357	2.67499
		25	102.381	1.64655
		12.5	95.5357	4.81779
		6.25	98.8095	2.57337
	3	100	0	0
		50	17.2205	3.335
		25	68.8822	5.39135
		12.5	95.1662	5.25463
		6.25	111.782	7.08676
	4	100	64.094	15.8075
		50	116.443	2.40601
		25	100	7.96939
		12.5	115.101	2.31181
		6.25	110.403	2.76687

Ceriodaphnia Survival				
Lab	Sample	Dilution	% Control	SD
A	1	100	100	0
		50	100	0
		25	100	0
		12.5	100	0
		6.25	100	0
	2	100	70	48.3046
		50	100	0
		25	100	0
		12.5	100	0
		6.25	100	0
	3	100	40	51.6398
		50	100	0
		25	100	0
		12.5	100	0
		6.25	100	0
	4	100	70	48.3046
		50	100	0
		25	100	0
		12.5	100	0
		6.25	100	0
B	1	100	80	42.1637
		50	90	31.6228
		25	100	0
		12.5	100	0
		6.25	100	0
	2	100	40	51.6398
		50	100	0
		25	90	31.6228
		12.5	100	0
		6.25	100	0
	3	100	100	0
		50	100	0
		25	100	0
		12.5	100	0
		6.25	100	0
	4	100	40	51.6398
		50	100	0
		25	100	0
		12.5	100	0

		6.25	100	0
C	1	100	90	31.6228
		50	80	42.1637
		25	100	0
		12.5	100	0
		6.25	90	31.6228
	2	100	0	0
		50	0	0
		25	40	51.6398
		12.5	100	0
		6.25	70	48.3046
	3	100	90	31.6228
		50	100	0
		25	100	0
		12.5	100	0
		6.25	100	0
	4	100	0	0
		50	0	0
		25	80	42.1637
		12.5	90	31.6228
		6.25	90	31.6228
E	1	100	100	0
		50	100	0
		25	100	0
		12.5	100	0
		6.25	100	0
	2	100	0	0
		50	90	31.6228
		25	100	0
		12.5	100	0
		6.25	100	0
	3	100	90	31.6228
		50	100	0
		25	100	0
		12.5	100	0
		6.25	100	0
	4	100	40	51.6398
		50	100	0
		25	90	31.6228
		12.5	100	0
		6.25	100	0

F	1	100	100	0
		50	100	0
		25	100	0
		12.5	100	0
		6.25	100	0
	2	100	0	0
		50	0	0
		25	0	0
		12.5	100	0
		6.25	100	0
	3	100	100	0
		50	100	0
		25	100	0
		12.5	100	0
		6.25	100	0
	4	100	0	0
		50	0	0
		25	0	0
		12.5	100	0
		6.25	100	0
G	1	100	100	0
		50	100	0
		25	100	0
		12.5	100	0
		6.25	100	0
	2	100	10	31.6228
		50	80	42.1637
		25	90	31.6228
		12.5	100	0
		6.25	100	0
	3	100	80	42.1637
		50	100	0
		25	100	0
		12.5	100	0
		6.25	100	0
	4	100	30	48.3046
		50	80	42.1637
		25	90	31.6228
		12.5	100	0
		6.25	100	0
H	1	100	100	0

		50	100	0
		25	90	31.6228
		12.5	100	0
		6.25	100	0
	2	100	90	31.6228
		50	100	0
		25	100	0
		12.5	90	31.6228
		6.25	100	0
	3	100	50	52.7046
		50	90	31.6228
		25	100	0
		12.5	90	31.6228
		6.25	100	0
	4	100	80	42.1637
		50	70	48.3046
		25	100	0
		12.5	90	31.6228
		6.25	90	31.6228
I	1	100	60	51.6398
		50	70	48.3046
		25	100	0
		12.5	100	0
		6.25	90	31.6228
	2	100	40	51.6398
		50	100	0
		25	100	0
		12.5	90	31.6228
		6.25	100	0
	3	100	100	0
		50	100	0
		25	100	0
		12.5	100	0
		6.25	100	0
	4	100	44.4444	51.6398
		50	111.111	0
		25	88.8889	42.1637
		12.5	111.111	0
		6.25	111.111	0

Hyalella Survival				
Lab	Sample	Dilution	% Control	SD
A	1	100	95.9184	5.47723
		50	102.041	0
		25	95.9184	5.47723
		12.5	102.041	0
		6.25	100	4.47214
	3	100	95.9184	8.94427
		50	102.041	0
		25	102.041	0
		12.5	102.041	0
		6.25	100	4.47214
B	1	100	92.36	7.73615
		50	98	4.47214
		25	100	0
		12.5	100	0
		6.25	100	0
	2	100	100	0
		50	100	0
		25	100	0
		12.5	100	0
		6.25	100	0
	3	100	94.54	8.13929
		50	100	0
		25	100	0
		12.5	98	4.47214
		6.25	100	0
	4	100	98	4.47214
		50	100	0
		25	100	0
		12.5	100	0
		6.25	100	0
C	2	100	0	0
		50	42.5	9.57427
		25	80	14.1421
		12.5	87.5	9.57427
		6.25	85	5.7735
	4	100	0	0
		50	47.3684	17.3205
		25	73.6842	14.1421
		12.5	92.1053	5

		6.25	100	5.7735
F	1	100	100	0
		50	100	0
		25	100	0
		12.5	100	0
		6.25	100	0
	2	100	45	10
		50	60	16.3299
		25	100	0
		12.5	100	0
		6.25	100	0
	3	100	80	0
		50	85	10
		25	100	0
		12.5	100	0
		6.25	100	0
	4	100	85	10
		50	100	0
		25	100	0
		12.5	100	0
		6.25	100	0
H	1	100	100	0
		50	100	0
		25	100	0
		12.5	100	0
		6.25	100	0
	2	100	45	10
		50	60	16.3299
		25	100	0
		12.5	100	0
		6.25	100	0
	3	100	80	0
		50	85	10
		25	100	0
		12.5	100	0
		6.25	100	0
	4	100	85	10
		50	100	0
		25	100	0
		12.5	100	0
		6.25	100	0

I	1	100	50	10
		50	105.556	10
		25	111.111	0
		12.5	105.556	10
		6.25	111.111	0
	2	100	20	16.3299
		50	95	10
		25	100	0
		12.5	90	20
		6.25	95	10
	3	100	100	10
		50	105.263	0
		25	105.263	0
		12.5	105.263	0
		6.25	105.263	0
	4	100	35	19.1485
		50	90	11.547
		25	80	16.3299
		12.5	95	10
		6.25	90	11.547
J	1	100	100	0
		50	100	0
		25	100	0
		12.5	100	0
		6.25	95	10
	2	100	10	11.547
		50	70	11.547
		25	75	10
		12.5	90	11.547
		6.25	100	0
	3	100	15	19.1485
		50	75	19.1485
		25	85	19.1485
		12.5	100	0
		6.25	100	0
	4	100	85	19.1485
		50	100	0
		25	90	11.547
		12.5	100	0
		6.25	95	10

Strongyloncentrotus % Normal				
Lab	Sample	Dilution	% Control	SD
A	1	100	98.1604	2.24151
		50	97.3537	1.58614
		25	85.8212	5.99931
		12.5	97.0162	2.5435
		6.25	94.4059	1.81314
	2	100	93.1099	3.61784
		50	80.5647	8.18751
		25	76.1982	7.56019
		12.5	77.157	6.15544
		6.25	90.732	2.39396
	3	100	0.000	0
		50	0.000	0
		25	0.000	0
		12.5	14.6478	2.35571
		6.25	24.6145	1.54593
	4	100	84.2386	2.55041
		50	90.0636	4.96291
		25	78.5829	6.74441
		12.5	89.9325	3.06065
		6.25	91.6255	4.64383
F	1	100	98.9362	2
		50	98.9362	2
		25	99.7872	1.30384
		12.5	98.5106	2.07364
		6.25	102.553	1.51658
	2	100	0	0
		50	2.34043	0.83666
		25	32.3404	5.54977
		12.5	98.7234	3.11448
		6.25	98.0851	1.92354
	3	100	68.2979	1.64317
		50	93.1915	3.20936
		25	97.8723	2
		12.5	98.0851	1.92354
		6.25	99.3617	1.34164
	4	100	0	0
		50	2.99145	1.48324
		25	32.265	1.92354
		12.5	88.0342	5.50454

		6.25	98.5043	1.78885
H	1	100	98.9362	2
		50	98.9362	2
		25	99.7872	1.30384
		12.5	98.5106	2.07364
		6.25	102.553	1.51658
	2	100	0	0
		50	2.34043	0.83666
		25	32.3404	5.54977
		12.5	98.7234	3.11448
		6.25	98.0851	1.92354
	3	100	68.2979	1.64317
		50	93.1915	3.20936
		25	97.8723	2
		12.5	98.0851	1.92354
		6.25	99.3617	1.34164
	4	100	0	0
		50	2.99145	1.48324
		25	32.265	1.92354
		12.5	88.0342	5.50454
		6.25	98.5043	1.78885
I	1	100	98.4416	1.70783
		50	99.7403	0.8165
		25	100	1.25831
		12.5	99.7403	1.82574
		6.25	99.2208	2.38048
	2	100	79.4038	6.70199
		50	93.2249	4.24264
		25	103.252	2.21736
		12.5	102.439	1.91485
		6.25	101.355	2.51661
	3	100	0	0
		50	0	0
		25	0	0
		12.5	0	0
		6.25	12.1813	1.5
	4	100	80.1061	3.87298
		50	93.634	1.70783
		25	99.2042	2.51661
		12.5	98.1432	1.73205
		6.25	99.7347	0.8165

J	1	100	100	2.44949
		50	97.5258	3.20936
		25	99.7938	1.92354
		12.5	101.649	1.14018
		6.25	101.237	0.83666
	2	100	36.6667	9.98499
		50	100.208	1.30384
		25	102.292	1.64317
		12.5	102.292	1.78885
		6.25	102.5	1.14018
	3	100	0	0
		50	0	0
		25	0	0
		12.5	0	0
		6.25	45.4918	9.88939
	4	100	62.7049	9.14877
		50	97.3361	0.70711
		25	98.9754	2.07364
		12.5	99.5902	2.94958
		6.25	100.82	1.14018

Mytilus % Normal-Alive				
Lab	Sample	Dilution	% Control	SD
E	1	100	98.2342	12.0082
		50	102.167	4.58662
		25	100.16	7.18713
		12.5	103.772	4.63435
		6.25	101.043	9.73531
	2	100	0	0
		50	13.4987	23.6468
		25	75.1734	11.1313
		12.5	94.1018	2.12015
		6.25	99.8447	5.81992
	3	100	0.74678	0.52581
		50	1.1619	0.52597
		25	5.64414	4.76283
		12.5	63.7351	18.3177
		6.25	99.5039	9.99032
	4	100	0.16853	0.31716
		50	5.30669	5.79639
		25	84.9918	7.55592

		12.5	99.917	7.44268
		6.25	100.912	4.36917
F	1	100	98.4848	1.58114
		50	99.3506	0.83666
		25	98.7013	1.30384
		12.5	100	1.14018
		6.25	98.7013	1.30384
	2	100	0	0
		50	0	0
		25	5.65217	1.92354
		12.5	68.4783	2.73861
		6.25	91.3043	5.70088
	3	100	93.8462	3.91152
		50	100.44	1.14018
		25	100.879	2.38747
		12.5	100	1.58114
		6.25	100.44	1.14018
	4	100	0	0
		50	0	0
		25	18.4211	3.76829
		12.5	81.7982	4.03733
		6.25	99.1228	2.07364
H	1	100	97.0894	2.07364
		50	98.3368	1.51658
		25	97.0894	1.14018
		12.5	100.832	1.22474
		6.25	98.3368	3.20936
	2	100	0	0
		50	2.89855	1.30384
		25	21.5321	4.60435
		12.5	76.6046	6.78233
		6.25	91.3043	4.02492
	3	100	81.3704	1
		50	91.4347	0.54772
		25	97.2163	5.71839
		12.5	100.857	3.27109
		6.25	101.285	1.51658
	4	100	0	0
		50	18.8912	5.81378
		25	60.1643	6.10737
		12.5	80.2875	3.89872

		6.25	92.4025	2.54951
I	1	100	102.023	4.98665
		50	93.9858	6.43661
		25	102.46	4.8833
		12.5	104.948	3.55375
		6.25	101.476	6.45446
	2	100	0	0
		50	0	0
		25	103.553	4.17852
		12.5	98.5128	4.21535
		6.25	100.881	7.59797
	3	100	0	0
		50	0	0
		25	75.9836	4.07063
		12.5	96.883	4.24028
		6.25	98.2626	3.52184
	4	100	0	0
		50	0.13537	0.2675
		25	83.0719	2.36414
		12.5	82.5152	11.769
		6.25	92.004	6.57368
J	1	100	99.4883	1.39498
		50	99.7473	1.8313
		25	97.7701	4.97951
		12.5	100.251	0.9464
		6.25	99.7094	2.35869
	2	100	0	0
		50	4.20139	0.92281
		25	95.6449	2.5347
		12.5	96.9736	4.70276
		6.25	92.5121	6.94413
	3	100	0	0
		50	0	0
		25	20.1938	4.62619
		12.5	93.8968	1.72533
		6.25	98.7495	3.5951
	4	100	0	0
		50	0.61011	0.554
		25	74.8547	6.43344
		12.5	99.5399	5.44517
		6.25	102.071	4.17242

ROUND 2 TABLES

Ceriodaphnia Reproduction				
Lab	Sample	Dilution	% Control	SD
A	1	100	69.5238	7.42668
		50	91.4286	5.09466
		25	76.1905	9.43987
		12.5	80.4762	3.41402
		6.25	72.381	5.30827
	2	100	0	0
		50	68.2081	8.08015
		25	103.468	8.60814
		12.5	100.578	9.20386
		6.25	85.5491	10.6333
	3	100	8.09249	3.27278
		50	52.6012	4.79467
		25	99.422	10.4009
		12.5	129.48	7.33636
		6.25	106.936	12.5277
	4	100	88.3721	2.85968
		50	100	6.0148
		25	111.628	3.85285
		12.5	112.209	6.07454
		6.25	101.744	5.33854
B	1	100	97.0803	2.50333
		50	89.781	2.71621
		25	94.1606	2.65832
		12.5	101.734	4.1897
		6.25	108.394	2.94581
	2	100	41.5033	7.81807
		50	77.1242	5.37897
		25	84.6405	1.91195
		12.5	84.3137	2.65832
		6.25	86.6013	3.50397
	3	100	10.5442	1.79196
		50	66.6667	2.27058
		25	83.7585	2.13391
		12.5	86.3946	2.17051
		6.25	95.994	1.85592
	4	100	88.6525	2.35702
		50	87.9433	5.09466
		25	91.844	2.80674

		12.5	97.5177	1.90029
		6.25	98.5816	1.61933
C	1	100	94.8498	12.7754
		50	95.279	11.5931
		25	80.2575	12.824
		12.5	91.8455	8.6564
		6.25	71.2446	11.8902
	2	100	46.3054	5.35828
		50	108.374	3.8873
		25	126.108	4.85798
		12.5	114.286	8.49575
		6.25	96.0591	8.70823
	3	100	52.0408	5.22388
		50	91.3265	4.8637
		25	106.633	8.5173
		12.5	107.143	4.6428
		6.25	120.918	3.335
	4	100	113.706	10.3086
		50	113.706	5.23238
		25	119.289	9.24061
		12.5	126.904	5.41603
		6.25	132.995	5.26624
F	1	100	49.2462	4.96208
		50	85.3403	5.2504
		25	109.424	8.17109
		12.5	106.283	8.55115
		6.25	69.1099	6.7297
	2	100	0	0
		50	0	0
		25	0	0
		12.5	0	0
		6.25	31.0526	7.35527
	3	100	0	0
		50	97.8142	7.29459
		25	123.497	10.5325
		12.5	146.448	12.2456
		6.25	112.022	9.77809
	4	100	76.8421	5.52167
		50	67.3684	5.3707
		25	58.4211	4.38305
		12.5	94.2105	9.8257

		6.25	54.2105	3.26769
I	1	100	57.4132	15.0096
		50	110.095	2.96086
		25	100	9.76445
		12.5	108.833	2.59272
		6.25	81.388	8.77876
	2	100	67.4267	10.2204
		50	94.4625	8.21922
		25	97.3941	7.53437
		12.5	115.961	2.41293
		6.25	105.863	2.99073
	3	100	8.36013	2.59058
		50	75.5627	2.59272
		25	100.322	3.58391
		12.5	105.145	4.24395
		6.25	109.325	6.21825
	4	100	35.2113	9.91351
		50	91.831	6.51835
		25	88.4507	8.4748
		12.5	98.5915	2.23607
		6.25	99.7183	2.71621
J	1	100	70.7792	8.25698
		50	83.4416	5.73585
		25	86.3636	2.79682
		12.5	90.2597	5.2451
		6.25	94.8052	5.65292
	2	100	90.4025	6.97296
		50	93.1889	3.92853
		25	102.167	5.57773
		12.5	109.288	2.26323
		6.25	106.192	3.335
	3	100	40.1294	6.13188
		50	94.822	4.16467
		25	87.7023	6.40226
		12.5	88.0259	4.1042
		6.25	92.5566	4.00555
	4	100	100.717	10.0714
		50	107.527	5.59762
		25	102.509	5.31664
		12.5	90.681	9.15363
		6.25	106.452	5.01221

Ceriodaphnia Survival				
Lab	Sample	Dilution	% Control	SD
A	1	100	80	42.1637
		50	100	0
		25	70	48.3046
		12.5	100	0
		6.25	100	0
	2	100	0	0
		50	77.7778	48.3046
		25	98.7654	33.3333
		12.5	88.8889	42.1637
		6.25	77.7778	48.3046
	3	100	88.8889	42.1637
		50	100	31.6228
		25	86.4198	44.0959
		12.5	111.111	0
		6.25	86.4198	44.0959
	4	100	100	31.6228
		50	100	31.6228
		25	100	31.6228
		12.5	88.8889	42.1637
		6.25	100	31.6228
B	1	100	111.111	0
		50	111.111	0
		25	111.111	0
		12.5	111.111	0
		6.25	111.111	0
	2	100	70	48.3046
		50	90	31.6228
		25	100	0
		12.5	100	0
		6.25	100	0
	3	100	100	0
		50	100	0
		25	100	0
		12.5	100	0
		6.25	100	0
	4	100	100	0
		50	90	31.6228
		25	100	0
		12.5	100	0

		6.25	100	0
C	1	100	88.8889	42.1637
		50	100	31.6228
		25	88.8889	42.1637
		12.5	100	31.6228
		6.25	88.8889	42.1637
	2	100	112.5	31.6228
		50	125	0
		25	125	0
		12.5	125	0
		6.25	112.5	31.6228
	3	100	90	31.6228
		50	100	0
		25	100	0
		12.5	100	0
		6.25	100	0
	4	100	125	0
		50	112.5	31.6228
		25	112.5	31.6228
		12.5	125	0
		6.25	125	0
F	1	100	100	31.6228
		50	100	31.6228
		25	111.111	0
		12.5	100	31.6228
		6.25	100	31.6228
	2	100	0	0
		50	0	0
		25	0	0
		12.5	0	0
		6.25	50	52.7046
	3	100	70	48.3046
		50	90	31.6228
		25	100	0
		12.5	100	0
		6.25	100	0
	4	100	88.8889	42.1637
		50	100	31.6228
		25	111.111	0
		12.5	100	31.6228
		6.25	111.111	0

I	1	100	77.7778	48.3046
		50	111.111	0
		25	111.111	0
		12.5	100	31.6228
		6.25	55.5556	52.7046
	2	100	77.7778	48.3046
		50	88.8889	42.1637
		25	88.8889	42.1637
		12.5	111.111	0
		6.25	111.111	0
	3	100	111.111	0
		50	111.111	0
		25	111.111	0
		12.5	111.111	0
		6.25	111.111	0
	4	100	90	31.6228
		50	100	0
		25	80	42.1637
		12.5	100	0
		6.25	100	0
J	1	100	100	0
		50	100	0
		25	100	0
		12.5	100	0
		6.25	100	0
	2	100	100	0
		50	100	0
		25	100	0
		12.5	100	0
		6.25	100	0
	3	100	90	31.6228
		50	100	0
		25	90	31.6228
		12.5	100	0
		6.25	100	0
	4	100	31.248	15.5964
		50	100	0
		25	100	0
		12.5	90	31.6228
		6.25	100	0

Hyalella Survival				
Lab	Sample	Dilution	% Control	SD
A	1	100	103.571	8.16497
		50	107.143	0
		25	103.571	8.16497
		12.5	107.143	0
		6.25	107.143	0
	2	100	86.2069	15.0555
		50	103.448	0
		25	100	8.16497
		12.5	103.448	0
		6.25	103.448	0
	3	100	80	12.6491
		50	90	10.9545
		25	100	0
		12.5	97.1667	6.94022
		6.25	100	0
	4	100	100	0
		50	100	0
		25	93.3333	10.328
		12.5	100	0
		6.25	96.6667	8.16497
B	1	100	100	0
		50	100	0
		25	100	0
		12.5	100	0
		6.25	100	0
	2	100	20	12.6491
		50	90	10.9545
		25	96.6667	8.16497
		12.5	100	0
		6.25	100	0
	3	100	70	20.9762
		50	96.6667	8.16497
		25	90	16.7332
		12.5	100	0
		6.25	100	0
	4	100	96.6667	8.16497
		50	100	0
		25	96.6667	8.16497
		12.5	100	0

		6.25	100	0
C	1	100	103.448	0
		50	100	8.16497
		25	103.448	0
		12.5	100	8.16497
		6.25	103.448	0
	2	100	56.6667	15.0555
		50	90	16.7332
		25	100	0
		12.5	100	0
		6.25	96.6667	8.16497
	3	100	65.5172	23.3809
		50	100	8.16497
		25	89.6552	16.3299
		12.5	103.448	0
		6.25	103.448	0
	4	100	107.407	8.16497
		50	111.111	0
		25	107.407	8.16497
		12.5	111.111	0
		6.25	111.111	0
F	1	100	100	0
		50	100	0
		25	100	0
		12.5	100	0
		6.25	95	10
	2	100	31.5789	11.547
		50	100	10
		25	94.7368	11.547
		12.5	105.263	0
		6.25	100	10
	3	100	90	20
		50	95	10
		25	100	0
		12.5	100	0
		6.25	95	10
	4	100	90	11.547
		50	85	19.1485
		25	100	0
		12.5	90	11.547
		6.25	95	10

I	1	100	100	0
		50	100	0
		25	93.3333	10.328
		12.5	100	0
		6.25	100	0
	2	100	50	20.6559
		50	100	10.328
		25	107.143	0
		12.5	107.143	0
		6.25	107.143	0
	3	100	80	12.6491
		50	96.6667	8.16497
		25	100	0
		12.5	100	0
		6.25	100	0
	4	100	96.6667	8.16497
		50	100	0
		25	100	0
		12.5	100	0
		6.25	100	0
J	1	100	96.6667	8.16497
		50	96.6667	8.16497
		25	100	0
		12.5	100	0
		6.25	96.6667	8.16497
	2	100	40	28.2843
		50	96.6667	8.16497
		25	93.3333	10.328
		12.5	100	0
		6.25	100	0
	3	100	86.6667	16.3299
		50	100	0
		25	90	10.9545
		12.5	100	0
		6.25	100	0
	4	100	96.6667	8.16497
		50	100	0
		25	93.3333	16.3299
		12.5	100	0
		6.25	96.6667	8.16497

Attachment 2

Development of Quality Assurance Guidance for the *Ceriodaphnia dubia* Reproduction Test

The accuracy and comparability of toxicity test information is vital to stormwater and effluent monitoring programs. The SMC has a primary goal of sharing data, and chief amongst the concerns for sharing data is comparability. Data quality is also important for wastewater effluent discharge monitoring programs, as findings of effluent toxicity may lead to costly studies to confirm the toxicity and identify the cause.

A recent toxicity intercalibration study by the SMC identified instances of poor comparability in the *Ceriodaphnia dubia* reproduction test. A sample of laboratory dilution water, prepared in accordance with standardized guidance and expected to be nontoxic, was found to be toxic by some of the participating labs. A repeat of the testing, conducted after modest standardization of test methods, also identified toxicity in the dilution water sample. Furthermore, the relative comparability of the laboratories varied between testing events. The SMC intercalibration study was unable identify the cause of the low comparability and recommended further investigation of this issue.

The goal of this project is to conduct studies to identify laboratory quality assurance practices that will improve comparability of the *C. dubia* reproduction test. The focus of this work is to identify test conditions and procedures that will minimize instances of toxicity when laboratory dilution water is tested by multiple laboratories. Improving the comparability of test results for samples expected to be nontoxic will improve confidence in the use and interpretation of the *C. dubia* reproduction test for effluent and stormwater monitoring. One of the outcomes of this study will be a guidance document that describes test procedures and quality assurance steps to improve toxicity data comparability.

A Technical Advisory Committee (TAC) composed of regulatory agencies (SWRCB, EPA, RWQCB), regulated agencies (stormwater, wastewater), and testing laboratories will be established to guide study design and review interpretation of the results. The proposed study design is composed of four major elements, all using the *C. dubia* reproduction test: 1) identification of factors influencing comparability, 2) optimization of test conditions, 3) confirmation testing, and 4) reporting. Identification of test factors in the first study element will be accomplished by conducting two rounds of testing by multiple laboratories. A set of different types of laboratory dilution water and positive controls will be tested by a large number of laboratories (10 or more) using EPA standard methods. Results from these tests will be used to identify a subset of factors, such as hardness, ionic composition, or culture feeding regime that appear to influence the results between labs.

The second study element (optimization of test conditions) will include targeted studies to identify specific combinations of test conditions that result in the best intra- and interlaboratory comparability of results. Testing will be conducted by a subset of the laboratories participating in the first study element. Results from this study element are expected to identify refinements to the *C. daphnia* test method that improve test comparability.

In the third element, the efficacy of the method refinements during the second element will be evaluated by conducting another round of testing of laboratory dilution water by a large number of laboratories. The results will be compared to those obtained in the first element to document improvements in data comparability.

The final element of the study will consist of preparing a report and database. The report will summarize the study results and describe procedure to improve test quality and comparability. A database containing raw data from the study will be made publicly available to facilitate public review and use of the information.

This project will require 36 months to complete and will be coordinated by SCCWRP. SCCWRP will be responsible for securing laboratories, establishing the TAC, creating the study design, coordinating sample preparation and testing, data evaluation, and reporting. Approximately 50% of project funding will be used to compensate testing laboratories for participating in the study.

Preliminary cost estimates for this project is $700,000. See Table below for costs associated by task.

TASK	**COST**
Identification of Factors Influencing Comparability	$300,000
Optimization of Test Conditions	$100,000
Confirmation Testing	$150,000
Advisory Committee, Data Management and Analysis, Reporting	$150,000
Total	**$700,000**