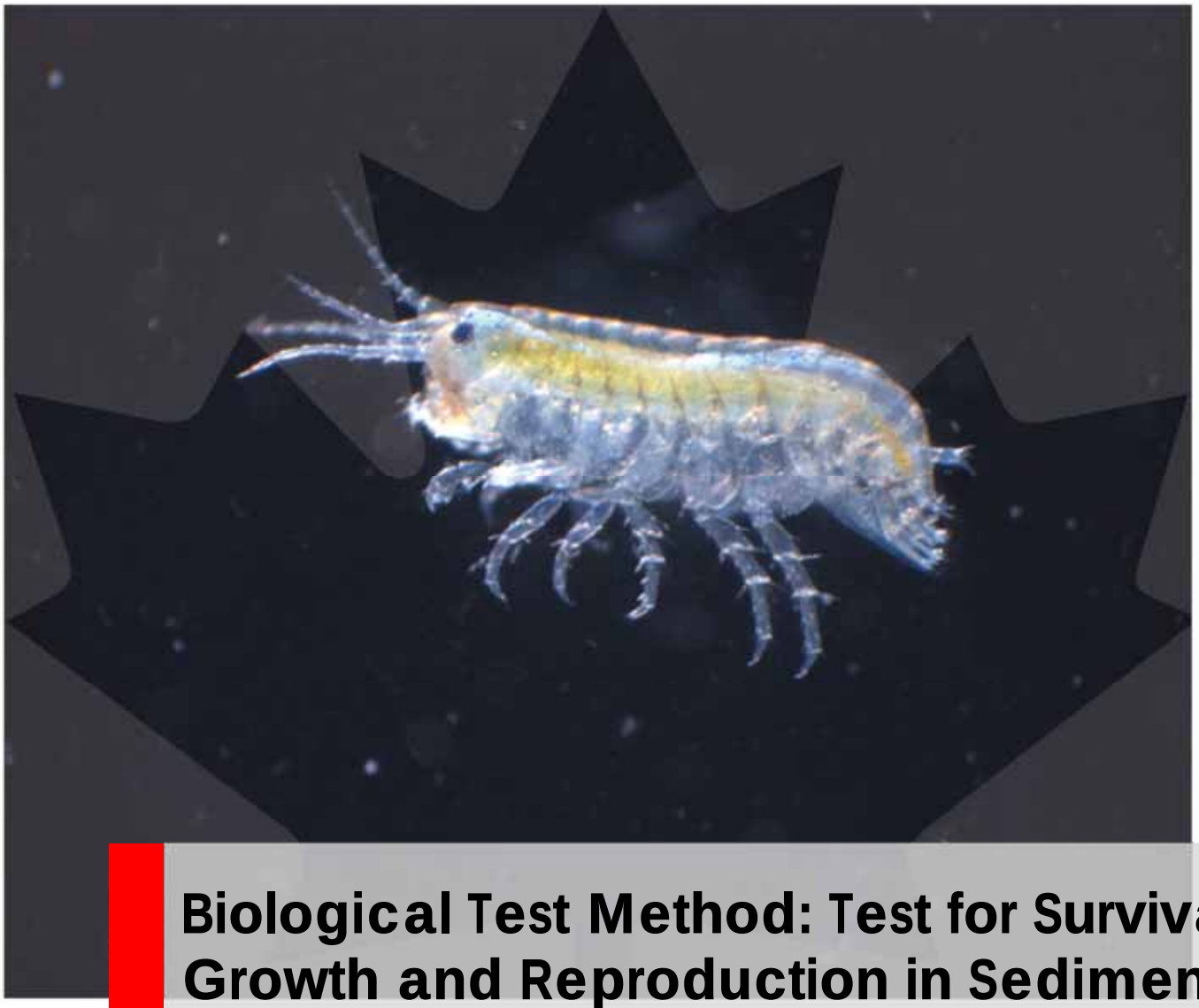
Environment and Climate Change Canada
Environnement et Changement climatique Canada
Canada



Biological Test Method: Test for Survival, Growth and Reproduction in Sediment and Water Using the Freshwater Amphipod *Hyalella azteca*

Report RM/33
Third Edition - September 2017

PDF:
Cat. No.: En49-7/1-33-2017E-PDF
ISBN: 978-0-660-09873-9

Paper:
Cat. No.: En49-7/1-33-2017E
ISBN: 978-0-660-09875-3

Environment and Climate Change Canada
Public Inquiries Centre
12th Floor, Fontaine Building
200 Sacré-Coeur Boulevard
Gatineau, QC K1A 0H3
Telephone: 819-938-3860
Toll Free: 1-800-668-6767 (in Canada only)
Email: ec.enviroinfo.ec@canada.ca

Photo on cover: Hyalella azteca © Environment and Climate Change Canada

Aussi disponible en français

Biological Test Method: Test for Survival, Growth and Reproduction in Sediment and Water Using the Freshwater Amphipod *Hyalella azteca*

Method Development and Applications Unit
Science and Technology Branch
Environment and Climate Change Canada
Ottawa, Ontario

Report RM/33
Third Edition - September 2017

Readers' Comments

Comments regarding the content of this report should be addressed to:

Richard Scroggins, Chief
Biological Assessment
and Standardization Section
Science and Technology Branch
Environment and Climate Change Canada
335 River Road
Ottawa ON K1A OH3

Review Notice

This report has been reviewed by the staff of the Environmental Technology Advancement Directorate, Environment and Climate Change Canada. Mention of trade names or commercial products does not constitute endorsement by Environment and Climate Change Canada for use. Other products of similar value are available.

Abstract

Revised methods now recommended by Environment and Climate Change Canada for performing sediment and water-only toxicity tests using the freshwater amphipod *Hyalella azteca* are described in this report. The endpoints for these tests are survival and dry weight of amphipods at the end of a 14-day test. This method also includes a reproduction endpoint over a longer term exposure (42 days). Key to successfully conducting the longer-term test is the recovery from sediment and accurate counting of the young of the adults at the end of the 42 day test, and as such, the 42-d methods include a performance criterion that requires laboratories to meet a specified % recovery. This revised version of Report RM/33 supersedes Environment and Climate Change Canada's test for survival and growth in sediment using *Hyalella azteca*, which was published as Report EPS 1/RM/33 in January 2013. The sediment test is intended primarily for measuring the adverse effect(s) of freshwater sediments, although procedures for testing estuarine sediments (≤15‰ salinity) are also described.

The 14-day survival and growth sediment toxicity test is normally conducted at 23 ± 1°C in glass beakers or jars containing a 100-mL layer of sediment and 175 mL of overlying water. An option for using a 1:4 sediment to water ratio is included herein for studies requiring greater volumes of overlying water for water-quality monitoring and/or chemical analyses. The test may be run as a single-concentration assay (e.g., for testing undiluted samples of field-collected sediment), or as a multi-concentration assay (e.g., for testing spiked-sediment or sediment mixtures at several concentrations) to determine the threshold of effect. For a single-concentration assay, a minimum of 5 replicate samples of sediment (i.e., field replicates) are collected at each discrete sampling station, and each one is tested for its toxicity to *H. azteca* as a single replicate. For a multi-concentration assay, a minimum of 5 replicate vessels (i.e., laboratory replicates) per treatment are required. Each replicate vessel contains 10 *H. azteca*. Amphipods are 2 to 9 days old and ranging in age by ≤3 days at the start of the test.

The 14-day survival and growth sediment test is routinely carried out as a static (i.e., no renewal) exposure, during which the overlying water is continuously aerated. If, however, the test water overlying sediment from any reference sampling station deteriorates or becomes fouled (i.e., due to high levels of ammonia, pH levels outside the tolerance range of *Hyalella azteca*, and/or low levels of dissolved oxygen) at any time during the test, and the objectives of the test are to assess toxic effects due to substances or materials without the deleterious or modifying effect of these confounding factors, the test must be carried out, or continued, as a static-renewal test. In the static-renewal exposure, the overlying water is renewed a minimum of 3 times weekly on non-consecutive days, at a rate of 2 volume additions in 24 hours. The animals are fed either a mixture of yeast, cereal grass media, and trout chow (YCT); ground commercial fish food flakes; or a combination of both YCT and fish food flakes. Food is added to each test vessel, either daily or 3 times per week on non-consecutive days. Selection of either feeding option depends on the objectives of the study and perhaps also on regulatory guidelines or requirements.

The 14-day water-only survival and growth test is conducted under the same exposure conditions as the sediment test, and shares many aspects of the test design. The test is carried out as a static-renewal exposure, with a minimum of 5 replicate vessels per treatment, each containing 275 mL of solution and a substrate. The water-only method has been included for use alone, or in conjunction

with the 14-day sediment test, which together might be useful in differentiating between historical contamination (i.e., from sediment) and current water and/or effluent quality.

The 42-day survival, growth and reproduction sediment toxicity test shares many aspects of the 14-day sediment test design, but with some notable differences. The 42-day test is conducted at 23 ± 2°C in 1 L glass beakers or jars containing a 18-mL layer of sediment and 900 mL of overlying water (for a sediment to water ratio of 1:50). The test can be conducted as single-concentration or a multi-concentration test. For a single-concentration test, a minimum of 8 replicate vessels (i.e., laboratory replicates) per treatment are tested for toxicity to *H. azteca*. For a multi-concentration test, a minimum of 5 replicate vessels (i.e., laboratory replicates) per treatment are required. Each replicate vessel contains 20 *H. azteca*. Amphipods must be 7 to 9 days old at the start of the test, and starting mean dry weight of the 7 to 9 day old amphipods should be 0.02 – 0.035 mg/individual.

The 42-d sediment test is conducted as a static-renewal exposure, during which the overlying water is continuously aerated. Overlying water is renewed on Days 14, 28 and 35. The test organisms are fed a mixture of yeast, cereal grass media and trout chow (YCT), and ground commercial fish food flakes. Food is added to each test vessel, 3 times per week on non-consecutive days. The amount of food added to each test vessel gradually increases during the 42-d exposure.

General or universal conditions and procedures are outlined for test preparation and performance. Additional conditions and procedures are stipulated that are specific to the intended use of the test. The sediment tests are suitable for measuring and assessing the toxicity of samples of field-collected sediment, sludge, or similar particulate material, or of sediment spiked (mixed) in the laboratory with chemical(s) or chemical substance(s), contaminated sediment, or other particulate material. The water-only test is suitable for measuring the toxicity of samples of industrial or sewage effluents, fresh waters (e.g., receiving water), aqueous extracts, or chemical substances. Instructions and requirements are included on test facilities, sample collection, handling and storing samples, culturing *H. azteca*, preparing sediment and aqueous mixtures, and initiating tests, specific test conditions, appropriate observations and measurements, endpoints and methods of calculation, and the use of reference toxicants.

Foreword

This is one of a series of **recommended methods** for measuring and assessing the toxic effect(s) on single species of aquatic or terrestrial organisms, caused by their exposure to samples of toxic or potentially toxic substances or materials under controlled and defined laboratory conditions. Recommended methods are those that have been evaluated by Environment and Climate Change Canada, and are favoured:

- for use in Environment and Climate Change Canada environmental toxicity laboratories;
- for testing that is contracted out by Environment and Climate Change Canada or requested from outside agencies or industry;
- in the absence of more specific instructions, such as are contained in regulations; and
- as a foundation for the provision of very explicit instructions as might be required in a regulatory protocol or standard reference method.

The different types of tests included in this series were selected because of their acceptability for the needs of environmental protection and management programs carried out by Environment and Climate Change Canada. These reports are intended to provide guidance and to facilitate the use of consistent, appropriate, and comprehensive procedures for obtaining data on the toxicity to aquatic or terrestrial life of samples of specific test substances or materials destined for or within the environment. Depending on the biological test method(s) chosen and the environmental compartment of concern, substances or materials to be tested for toxicity could include samples of chemical or chemical product, effluent, elutriate, leachate, receiving water, sediment or similar particulate material, or soil or similar particulate material. Appendix A provides a listing of the biological test methods and supporting guidance documents published to date by Environment and Climate Change Canada as part of this series.

Words defined in the Terminology section of this document are italicized when first used in the body of the report according to the definition. Italics are also used as emphasis for these and other words, throughout the report.

Table of Contents

List of Tables

List of Figures

List of Abbreviations and Chemical Formulae

ANOVA	analysis of variance
AVS	acid volatile sulphide
°C	degree(s) Celsius
$CaCO_3$	calcium carbonate
CdCl2	cadmium chloride
$CuSO_4$	copper sulphate
cm	centimetre(s)
CV	coefficient of variation
d	day(s)
DO	dissolved oxygen (concentration)
EC50	median effect concentration
Eh	reduction potential
g	gram(s)
g/kg	gram(s) per kilogram
h	hour(s)
HCl	hydrochloric acid
HCO_3	bicarbonate
ICp	inhibiting concentration for a (specified) percent effect
KCl	potassium chloride
kg	kilogram(s)
L	litre(s)
LC50	median lethal concentration
LOEC	lowest-observed-effect concentration
mg	milligram(s)
$MgSO_4$	magnesium sulphate
min	minute(s)
mL	millilitre(s)
mS	millisiemens
mm	millimetre(s)
NaBr	sodium bromide
NaCl	sodium chloride
$NaHCO_3$	sodium bicarbonate
NaOH	sodium hydroxide
NH_3	ammonia
nm	nanometre(s)
NOEC	no-observed-effect concentration
O_2	oxygen
SD	standard deviation
s	second(s)
spp.	species (plural)
SEM	simultaneously extracted metals
v:v	volume-to-volume
YCT	yeast, cereal grass media, and trout chow
YPF	young per surviving female
µE	microeinstein(s)
µg	microgram(s)
µm	micrometre(s)
µmhos/cm	micromhos per centimetre
TM (TM)	Trade Mark
>	greater than
<	less than
≥	greater than or equal to
≤	less than or equal to
/	per; alternatively, "or" (e.g., control/dilution water)
~	approximately
±	plus or minus
%	percentage or percent
‰	parts per thousand

Terminology

Note: All definitions are given in the context of the procedures in this report, and might not be appropriate in another context.

Grammatical Terms

Must is used to express an absolute requirement.

Should is used to state that the specified condition or procedure is recommended and ought to be met if possible.

May is used to mean "is (are) allowed to."

Can is used to mean "is (are) able to."

Might is used to express the possibility that something could exist or happen.

General Technical Terms

Acclimation is physiological adjustment to a particular level of one or more environmental factors such as temperature. The term usually refers to the adjustment to controlled laboratory conditions.

Ammonia means total ammonia [$NH_3 + NH_4$, as nitrogen (N)], un-ionized ammonia (NH_3, as N) and ionized ammonia (NH_4^+, as N). The percentage of un-ionized ammonia (NH_3) in total ammonia is determined by pH and temperature. The following formulae are used to calculate the fraction of un-ionized (NH_3) and ionized (NH_4^+) ammonia. Since $NH_3 = 1/(1 + 10^{pK - pH})$ and $NH_4^+ = 1/(1 + 10^{pH - pK})$, and total ammonia = $NH_3 + NH_4^+$, the concentration of un-ionized ammonia (assuming a pK of 9.56 at 15 ºC) is calculated as: un-ionized ammonia = (total ammonia) × $[1/(1 + 10^{pK - pH})]$ (USEPA, 1999b).

Batch means a single group of amphipods (e.g., 2- to 9-days old and ranging in age by ≤3 days for the 14-d test, or 7- to 9-days old for the 42-d exposure) taken from a *culture* at a discrete time, in order to provide all of the test organisms intended for use in a discrete toxicity test (including any associated reference toxicity test).

Compliance means in accordance with governmental regulations or requirements for issuing a permit.

Conductivity is a numerical expression of the ability of an aqueous solution to carry an electric current. This ability depends on the concentrations of ions in solution, their valence and mobility, and on the solution's temperature. Conductivity is reported (typically at a temperature of 25 °C) as micromhos per centimetre (μmhos/cm) or as millisiemens per metre (mS/m); 1 mS/m = 10 μmhos/cm.

Culture, as a noun, means the stock of animals or plants that is raised under defined and controlled conditions through one or more generations, to produce healthy test organisms. As a verb, it means to carry out the procedure of raising healthy test organisms from one or more generations, under defined and controlled conditions.

Flocculation is the formation of a light, loose precipitate (i.e., a floc) from a solution.

Flow-through describes apparatus or tests in which solutions or overlying water in *culture* or test vessels are/is renewed continuously by the constant inflow of a fresh solution.

Intermittent renewal describes a toxicity test in which test solutions or overlying water are/is renewed periodically during the test. Synonymous terms are *static renewal*, "batch replacement," "renewed static," "renewal," "static replacement," and "semistatic."

Lux is a unit of illumination based on units per square metre. One lux = 0.0929 foot-candles and one foot-candle = 10.76 lux. For conversion of lux to quantal flux [μmol/(m^2 · s)], the spectral quality of the light source must be known. Light conditions or irradiance are properly described in terms of quantal flux (photon fluence rate) in the photosynthetically effective wavelength range of approximately 400 to 700 nm. The relationship between quantal flux and lux or foot-candles is highly variable and depends on the light source, the light meter used, the geometrical arrangement, and the possibilities of reflections (see ASTM, 1999). An approximate conversion between quantal flux and lux, for full-spectrum fluorescent light (e.g., Vita-Lite® by Duro-Test®), is as follows: one lux is approximately equal to 0.016 μmol/(m^2 · s) (Deitzer, 1994; Sager and McFarlane, 1997).

Monitoring is the routine (e.g., daily, weekly, monthly, quarterly) checking of quality, or collection and reporting of information. In the context of this report, it means either the periodic (routine) checking and measurement of certain biological or water quality variables, or the collection and testing of samples of sediment, wastewater, or receiving water for toxicity.

Percentage (%) is a concentration expressed in parts per hundred. With respect to test substances or materials, 10 percent (10%) represents 10 units or parts of substance (or material) diluted with sediment or water to a total of 100 parts. Depending on the test substance or material, concentrations can be prepared on a weight-to-weight, weight-to-volume, or volume-to-volume basis, and are expressed as the percentage of test substance or material in the final sediment mixture or solution.

pH is the negative logarithm of the activity of hydrogen ions in gram equivalents per litre. The pH value expresses the degree or intensity of both acidic and alkaline reactions on a scale from 0 to

14, with 7 representing neutrality, numbers less than 7 indicating increasingly greater acidic reactions, and numbers greater than 7 indicating increasingly basic or alkaline reactions.

Photoperiod is the duration of illumination and darkness within a 24-h day.

Precipitation means the formation of a solid (i.e., precipitate) from some or all of the dissolved components of a solution.

Pretreatment means treatment of a sediment or water sample, or portion thereof, before exposure of amphipods.

Protocol is an explicit set of procedures for a test, formally agreed upon by the parties involved, and described precisely in a written document.

Reference method refers to a specific *protocol* for performing a toxicity test, i.e., a biological test method with an explicit set of test procedures and conditions, formally agreed upon by the parties involved and described precisely in a written document. Unlike other multi-purpose (generic) biological test methods published by Environment and Climate Change Canada, the use of a reference method is frequently restricted to testing requirements associated with specific regulations.

Salinity is the total amount of solid substance, in grams, dissolved in 1 kg of (sea)water; and is traditionally expressed as parts per thousand (‰). It is determined after all carbonates have been converted to oxides, all bromide and iodide have been replaced by chloride, and all organic matter has been oxidized. Salinity can also be measured directly using a salinity/conductivity meter or other means (see APHA *et al.*, 1995, 2005).

Static describes toxicity tests in which test solutions or overlying water are not renewed during the test.

Static-renewal – see *Intermittent-renewal*.

Turbidity is the extent to which the clarity of water has been reduced by the presence of suspended or other matter that causes light to be scattered and absorbed rather than transmitted in straight lines through the sample. It is generally expressed in terms of Mephelometric Turbidity Units.

Water renewal describes the renewal of the overlying water or test solutions in test vessels, on a regular and timed basis (e.g., three times weekly) throughout the test. This *may* be done manually or using an automated system that enables *intermittent renewal* of overlying water at a fixed rate.

Terms for Test Materials or Substances

Acid-Volatile Sulfide (AVS) is sulfide liberated (volatilized) from sediment by treatment with hydrochloric acid. See related term Simultaneously Extracted Metal (SEM).

Chemical is any element, compound, formulation, or mixture of a substance that might be mixed with, deposited in, or found in association with sediment or water, or enter the environment through spillage, application, or discharge.

Clean sediment is sediment that does not contain concentrations of any substance(s) causing discernible distress to the test organisms or reducing their survival or growth during the test.

Contaminated sediment is sediment containing chemical substances at concentrations that pose a known or potential threat to environmental or human health.

Control is a treatment in an investigation or study that duplicates all the conditions and factors that might affect the results of the investigation, except the specific condition being studied. In toxicity tests, the control must duplicate all the conditions of the exposure treatment(s), but must contain no contaminated test material or substance. The control is used as a check for the absence of measurable toxicity due to basic test conditions (e.g., quality of dilution water, health of test organisms, or effects due to their handling).

Control/dilution water for the water-only test, means the water used for diluting the test material or substance, or for the test *control* or both. For the sediment test it is the water used for preparing a series of concentrations of a test chemical, or that used as overlying water. Control/dilution water is frequently identical to the *culture* and *test (overlying) water*.

Control sediment is *clean* sediment not containing concentrations of one or more contaminants that could affect the survival, growth, or behaviour of the test organisms. Control sediment might be natural sediment from an uncontaminated site, or formulated (reconstituted) sediment. This sediment must contain no added test material or substance, and must enable acceptable (i.e., ≥80%) survival of the test organisms during the test. The use of control sediment provides a basis for interpreting data derived from toxicity tests using test sediment(s), and also provides a base sediment for spiking procedures.

Dechlorinated water is a chlorinated water (usually municipal drinking water) that has been treated to remove chlorine and chlorinated compounds from solution.

Deionized water is water that has been purified by passing it through resin columns or a reverse osmosis system.

Dilution water is the water used to dilute a test substance or material in order to prepare different concentrations for the various toxicity test treatments. (See also *control/dilution water* and *test water.)*

Distilled water is water that has been passed through a distillation apparatus of borosilicate glass or other material, to remove impurities.

Effluent is any liquid waste (e.g., industrial, municipal) discharged to the aquatic environment.

Elutriate is an aqueous solution obtained after adding water to a solid material (e.g., sediment, tailings, drilling mud, dredge spoil), shaking the mixture, then centrifuging or filtering it or decanting the supernatant.

Leachate is water or wastewater that has percolated through a column of soil or solid waste within the environment.

Material is the *substance* or substances from which something is made. A material would have more or less uniform characteristics. Sediment, effluent, leachate, elutriate, or surface water are materials. Usually, the material would contain several or many substances.

Overlying water is water placed over sediment in a test vessel. (See also *test water*.)

Pore water (also called *interstitial* water) is the water occupying space between sediment particles.

Receiving water is surface water (e.g., in a stream, river, or lake) that has received a discharged waste, or else is about to receive such a waste. Further descriptive information must be provided to indicate which meaning is intended.

Reconstituted water is high purity deionized or glass distilled water to which reagent grade chemicals have been added. The resultant synthetic fresh water should be free from contaminants and have the desired *pH*, alkalinity, and hardness characteristics. Reconstituted water can also be fresh water to which commercially available dry ocean salts, reagent-grade salts, or brine has been added, in a quantity that provides the seawater *salinity* (and pH) desired for culturing organisms and for testing purposes (e.g., for a test using estuarine sediment).

Redox potential (also known as oxidation-reduction potential) is a measure (in volts) of the affinity of a substance for electrons relative to hydrogen which is set at zero. Substances more strongly electronegative than hydrogen have positive redox potentials. Substances less electronegative than hydrogen have negative redox potentials. Redox potential affects metal speciation, thus affecting metal bioavailability. This measurement can be difficult to determine as Redox potential is influenced by the oxidative state of the sediment.

Reference sediment is a field-collected sample of presumably *clean* sediment, selected for properties (e.g., particle size, compactness, total organic content) representing sediment conditions that closely match those of the sample(s) of test sediment except for the degree of chemical contaminants. It is often selected from a site uninfluenced by the source(s) of contamination (i.e., reference site) but within the general vicinity of the sites where samples of test sediment are collected. (See also *site*.)

Reference toxicant is a standard chemical used to measure the sensitivity of the test organisms in order to establish confidence in the toxicity data obtained for a test material or substance. In most instances, a toxicity test with a reference toxicant is performed to assess the sensitivity of the

organisms at the time the test material or substance is evaluated, and the precision and reliability of results obtained by the laboratory for that chemical.

Reference toxicity test is a test conducted using a reference toxicant in conjunction with a sediment or water-only toxicity test, to appraise the sensitivity of the organisms at the time the test material or substance is evaluated, and the precision and reliability of results obtained by the laboratory for that chemical. Deviations outside an established normal range indicate that the sensitivity of the test organisms, and the performance and precision of the test, are suspect. A reference toxicity test is most often performed in the absence of sediment (i.e., as a *water-only* test), although it can also be conducted as a *spiked sediment* test.

Sampling station means a specific location, within a site where the sample(s) of field-collected sediment are obtained for toxicity tests and associated physicochemical analyses (see Figure 2). A reference sampling station is a specific location within a reference *site* where the *reference sediment* samples are collected. (See also *site*.)

Sediment is natural particulate material, which has been transported and deposited in water and usually lies below water. The term can also describe a substrate that has been experimentally prepared (formulated) using selected particulate material (sand of particular grain size, bentonite clay, etc.) and within which the test organisms can burrow.

Simultaneously Extracted Metals (SEM) involves treating metal-containing sediment with hydrochloric acid to liberate metal ions into the acid solution that were previously bound with acid volatile sulfides (plus any unbound metals). They are termed 'simultaneously extracted' because they are the metals that are liberated from sediment while the volatilization of sulfide is occurring (see term Acid Volatile Sulfide (AVS)). The ratio of SEM/AVS can be calculated and used to determine if metals are bioavailable.

Site means a delineated "tract" of sediment that is being used or considered as a study area, usually from the perspective of its being contaminated or potentially contaminated by human activity. A *reference site* is a site uninfluenced by the source(s) of contamination but within the general vicinity of the sites where samples of test sediment are collected (see Figure 2). (See also *reference sediment*.)

Solid-phase sediment (also called *whole* sediment) is the intact sediment used to expose the test organisms, not a form or derivative of the sediment such as pore water or a resuspended sediment.

Solvent control sediment is a sample of sediment included in a test involving *spiked sediment*, in which an organic solvent is required to solubilize the test chemical before mixing it in a measured quantity of *control sediment*. The amount of solvent used when preparing the solvent control sediment must contain the same concentration of solubilising agent as that present in the highest concentration of the test chemical(s) in the sample of spiked sediment to be tested. This concentration of solvent should not adversely affect the performance of *Hyalella* during the test. Any test that uses an organic solvent when preparing one or more concentrations of chemical-spiked sediment must include a solvent control sediment in the test. (See also *control sediment*, *chemical* and *spiked sediment*.)

Spiked sediment is any sediment (clean or contaminated) to which a test substance or material such as a chemical, a mixture of chemicals, drilling mud, contaminated dredge spoil, sludge, or contaminated sediment has been added experimentally and mixed thoroughly to evenly distribute the substance or material throughout the sediment.

Stock solution is a concentrated solution of the substance or material to be tested. Measured volumes of a stock solution are added to *dilution water* to prepare the required strengths of test solutions.

Substance is a particular kind of material having more or less uniform properties. The word *substance* has a narrower scope than *material*, and might refer to a particular chemical (e.g., an element) or chemical product.

Test sediment is a field-collected sample of solid-phase sediment, taken from a site thought to be contaminated with one or more chemicals, and intended for use in the toxicity test with amphipods. In some instances, the term also applies to any sediment sample or mixture of spiked sediment (including control and reference sediment) used in the test.

Test water is the water placed over the layer of sediment in the test vessels, i.e., *overlying* water. It also denotes the water used to manipulate the sediment, if necessary (e.g., for preparing formulated sediment or mixtures of spiked sediment, or for wet sieving), and that used as control/dilution water for water-only tests. (See also *control/dilution water*.)

Upstream water is surface water (e.g., in a stream, river, or lake) that is not influenced by the effluent (or other test material or substance), by virtue of being removed from it in a direction against the current or sufficiently far across the current.

Wastewater is a general term that includes effluents, leachates, and elutriates.

Statistical and Toxicological Terms

a priori literally refers to something that is independent of experience. In the context of test design and statistics, a priori tests are ones that have been planned before the data were collected. Test objectives and test design would influence the decisions as to which a priori tests to select. (See also *post hoc*.)

Biomass means the total (dry) weight of living *Hyalella* in a *replicate* at the end of the test, divided by the number of juveniles that started in the replicate. The biomass *endpoint* represents a combination of sublethal effect and mortality.

Coefficient of variation (CV) is the standard deviation (SD) of a set of data divided by the mean, expressed as a *percentage*. It is calculated according to the following formula:

$$CV\ (\%) = 100\ SD \div mean.$$

Continuous (variable) can take on any whole or fractional number on a numerical scale. The word *continuous* is synonymous with *quantitative* and is more commonly used by statisticians concerned with toxicology, especially Europe. (See also *quantitative*.)

Contrast analysis is used to compare the mortality response at different sampling stations. Mathematically, it involves the partitioning of the sums of squares of a categorical independent variable to test an *a priori* hypothesis as defined by a contrast.

Endpoint means the measurement(s) or value(s) that characterize the results of the test (e.g., LC50, IC25). It also means the response of the test organisms that is measured (e.g., death or increased weight of live organisms).

Geometric mean is the mean of repeated measurements, calculated on a logarithmic basis. It has the advantage that extreme values do not have as great an influence on the mean as is the case for an arithmetic mean. The *geometric mean* can be calculated as the n^{th} root of the product of the "n" values, and it can also be calculated as the antilogarithm of the mean of the logarithms of the "n" values.

Homoscedasticity refers herein to data showing homogeneity of the residuals within a scatter plot. This term applies when the variability of the residuals does not change significantly with that of the independent variable (i.e., the test concentrations or treatment levels). When performing statistical analyses and assessing residuals (e.g., using Levene's test), for test data demonstrating homoscedasticity (i.e., homogeneity of residuals), there is no significant difference in the variance of residuals across concentrations or treatment levels.

Hormesis is an effect in which low concentrations of the test material or substance act as a stimulant for performance of the test organisms compared to that for the control organisms (i.e., performance in one or more low concentrations is enhanced and "better" than that in the *control* treatment). At higher concentrations, deleterious effects are seen.

ICp is the inhibiting concentration for a (specified) percent effect. It represents a point estimate of the concentration of test material or substance that causes a designated percent impairment in a quantitative biological function such as growth. For example, an IC25 could be the concentration estimated to cause a 25% reduction in dry weight attained at the end of the test by the test organisms, relative to that in the control. This term should be used for any toxicological test that measures a continuously variable effect, such as dry weight at test end, reproduction, or respiration.

LC50 is the median lethal concentration, i.e., the concentration of substance or material in sediment (e.g., mg/kg) or water (e.g., mg/L) estimated to be lethal to 50% of the test organisms. The LC50 and its 95% confidence limits are usually derived by statistical analysis of mortalities in five or more test concentrations, after a fixed period of exposure. The duration of exposure must be specified (e.g., 96-h LC50 for a water-only reference toxicity test, or 14-d LC50 for a *survival-and-growth* toxicity test, using *Hyalella azteca*). Depending on the study objectives, an LCp other than LC50 (e.g., an LC25) might be calculated instead of or in addition to the LC50.

Lethal means causing death by direct action. Death of amphipods is defined as the cessation of all visible signs of movement or activity indicating life (e.g., absence of a pleopod twitch).

LOEC is the lowest-observed-effect concentration. This is the lowest concentration of a test substance or material to which organisms are exposed, that causes observed and statistically significant sublethal effects on the organism. For example, the LOEC might be the lowest concentration at which the dry weight of exposed organisms at test end was significantly less than that in the control groups.

Logistic regression, like all types of regression, investigates the relationship between a response, or dependent variable, and one or more independent variables. The specific features of the logistic regression used in this test method include: a binary response variable (mortality), and three different classes of independent variables (*continuous*, *ordinal*, and categorical). It is a linear model, and linearity is due to the logit transformation of the dependent variable.

Monotonic treatment-response, in the response variable, refers to the property of consistently increasing (or decreasing) over the range of the independent variable. A typical lethality curve is monotonic, because lethality increases as dose increases. In contrast, inhibition curves, which show *hormesis*, are non-monotonic, because there is low-dose stimulation, followed by inhibition, as dose increases.

NOEC is the no-observed-effect concentration. This is the highest concentration of a test substance or material to which organisms are exposed, that does not cause any observed and statistically significant sublethal effects on the organism. For example, the NOEC might be the highest test concentration at which an observed variable such as dry weight at test end is not decreased significantly from weight in the control groups.

Normality (or *normal distribution*) refers to a symmetric, bell-shaped array of observations. The array relates frequency of occurrence to the magnitude of the item being measured. In a *normal distribution*, most observations will cluster near the mean value, with progressively fewer observations toward the extremes of the range of values. The normal distribution plays a central role in statistical theory because of its mathematical properties. It is also central in biological sciences because many biological phenomena follow the same pattern. Many statistical tests assume that data are normally distributed, and therefore it can be necessary to test whether that is true for a given set of data.

Ordinal (variable) is a class of discrete data where there is a relative magnitude from low to high (e.g., no effect, minimal effect, high effect). In the context of this test method, "ordinal" is used to describe sampling stations which are expected to occur along a concentration gradient. That is, concentrations of the substances under study would be expected to sequentially increase or decrease along the specific ordering of the sampling stations. A common example would be sampling stations located downstream of a point source effluent. Also known as *ordered*, an *ordinal variable* is a variable which possesses the property of being *ordered*. (See also *unordered*.)

post-hoc literally refers to something performed after-the fact, or "after this." In the context of test design and statistics, *post-hoc* tests are those that are decided on after the data has been collected. Used in a more general sense, the purpose of the *post-hoc* test is to determine which treatment means are different from each other, while adjusting for the overall Type I error rate. (See also *a priori*.)

Precision refers to the closeness of repeated measurements of the same quantity to each other, i.e., the degree to which data generated from replicate measurements differ. It describes the degree of certainty around a result, or the tightness of a statistically derived endpoint such as an ICp.

Quantal is an adjective, as in quantal data, quantal test, etc. A quantal effect is one for which each test organism either shows the effect of interest or does not show it. For example, an animal might either live or die, or it might develop normally or abnormally. Quantal effects are typically expressed as numerical counts or percentages thereof.

Quantitative is an adjective, as in quantitative data, quantitative test, etc. A quantitative effect is one in which the measured effect can take any whole or fractional value on a numerical scale. An example would be the weight attained by individual organisms, or the number of progeny produced at the end of a test.

Replicate (test vessel) refers to a single test vessel containing a prescribed number of organisms in either one concentration or replicate sample of the test material or substance, or in the control or reference treatment(s). A replicate in a treatment must be an independent test unit; therefore, any transfer of organisms or test substance or material from one test vessel to another would invalidate a statistical analysis based on replication. The term is also used to refer to subsamples (i.e., laboratory replicates) of control sediment (Section 3.5), spiked sediment (Section 6.2), or water (Section 7.3), each of which is prepared in the laboratory. For control sediment, and for each treatment in the case of a multi-concentration test, there would normally be a minimum of five test vessels or replicates.

Replicate sample(s) are field-replicated samples collected from the same sampling station, to provide an estimate of the sampling error or to improve the precision of estimation (see Figure 2). A single sediment sample from a sampling station is treated as one replicate. Additional samples are considered to be additional replicate samples when they are treated identically but stored in separate sample containers (i.e., not composited).

Sublethal (*toxicity*) means detrimental to the organism but below the concentration or level of contamination that directly causes death within the test period.

Sublethal effect is an adverse effect on an organism below the concentration or level of contamination which directly causes death within the test period.

Survival-normalized reproduction reflects the fact that overall recruitment is influenced by both individual fecundity (i.e., number of offspring, expressed as young per female) and the number of adults surviving to reproduce. Survival-normalized reproduction is calculated by multiplying the young per female for a replicate by the fraction survival for that replicate.

Toxic means poisonous. A toxic chemical or material can cause adverse effects on living organisms if present in sufficient amount at the right location. *Toxic* is an adjective or adverb and should not be used as a noun; whereas *toxicant* is a legitimate noun.

Toxicant is a toxic substance or material.

Toxicity is the inherent potential or capacity of a substance or material to cause adverse effect(s) on living organisms. These effects could be *lethal* or *sublethal*.

Toxicity test is a procedure for determining the effect of a substance or material on a group of selected organisms (e.g., *H. azteca*), under defined conditions. An aquatic or sediment toxicity test usually measures: (a) the proportions of organisms affected (*quantal*); and/or (b) the degree of effect shown (*quantitative* or graded), after exposure to a specific test substance or material (e.g., a sample of sediment or *wastewater*) or mixture thereof (e.g., a chemical/sediment or chemical/water mixture).

Treatment is, in general, an intervention or procedure whose effect is to be measured. More specifically, in toxicity testing, it is a condition or procedure applied to the test organisms by an investigator with the intention of measuring the effects on those organisms. The treatment could be a specific concentration of a potentially toxic material or substance. Alternatively, a treatment might be a particular test material (e.g., a particular sample of contaminated, control, or reference sediment, as well as effluent, elutriate, leachate, receiving water, or control water).

Unordered is the absence of a gradient when referring to field sampling stations or independent variables. Also known as categorical, a common example would be field sampling stations located in a lake (see Figure 2). (See also *ordinal*.)

Warning chart is a graph used to follow changes over time in the endpoints for a reference toxicant. The date of the test is on the horizontal axis and the effect-concentration is plotted on the vertical logarithmic scale.

Warning limit is plus or minus two standard deviations, calculated on a logarithmic basis, from the historic geometric mean of the endpoints from toxicity tests with a reference toxicant.

Acknowledgements

The first edition of this biological test method, published in December 1997, was co-authored by D.J. McLeay (McLeay Environmental Ltd., Victoria, BC) and K.E. Day (Aquatic Ecosystem Conservation Branch, National Water Research Institute, Environment and Climate Change Canada, Burlington, ON).

With assistance from K. Conlan and E. Hendrycks of the Canadian Museum of Nature (Ottawa, ON), E.L. Bousfield (Research Associate, Royal British Columbia Museum, Victoria, BC) prepared the taxonomic and life history information in Appendix G. Testing support and method validation were provided by P. Hamr, S. Kirby, and P. Gillis within Environment and Climate Change Canada's laboratory facilities at the National Water Research Institute, Canada Centre for Inland Waters (Burlington, ON). Statistical guidance for the first edition of this method was written or reviewed and improved by J.B. Sprague (Sprague Associates Ltd., Salt Spring Island, BC).

The first edition of this report was based on existing reports (published or otherwise) that described procedures and conditions used in the United States and Canada for culturing *Hyalella azteca* and for conducting sediment toxicity tests using this species of freshwater amphipod. In particular, the specific procedures and conditions for culturing and testing *H. azteca* recommended by the United States Environmental Protection Agency (USEPA, 1994a), together with subsequent intra- and interlaboratory studies with Chironomus riparius undertaken by Canadian and U.S. researchers (Milani et al., 1996), guided the development of the first edition of this biological test method. With regard to the method development and validation studies (Milani et al., 1996), we gratefully acknowledge the participation and contribution of the following individuals and organizations: D. Milani and S. Kirby (National Water Research Institute, Environment and Climate Change Canada, Burlington, ON); K. Doe (Environment and Climate Change Canada, Moncton, NB); K. Holtze (B.A.R. Environmental Inc., Guelph, ON); A. Putt (Springborn Laboratories Inc., Wareham, MA); P. Riebel (Beak Consultants Ltd., Dorval, PQ); and G. van Aggelen (Environment and Climate Change Canada, North Vancouver, BC).

R.P. Scroggins (Method Development and Application Section, Environmental Technology Centre, Environment and Climate Change Canada, Gloucester, ON) acted as Scientific Authority for the first edition of this method and provided technical input and direction throughout the work. Members of the Inter-Governmental Ecotoxicological Testing Group (IGETG, Appendix B) participated in the development and review of the original report and are thanked accordingly. Special acknowledgement is made of the many useful comments provided by each member of the Environment and Climate Change Canada committee of scientific experts responsible for the initial and final reviews of the first edition of this report: G.T. Ankley (USEPA, Duluth, MN); U. Borgmann (Fisheries and Oceans Canada, Burlington, ON); G.A. Burton, Jr. (Wright State University, Dayton, OH); C.W. Hickey (NIWA Ecosystems, Hamilton, New Zealand); C.G. Ingersoll (U.S. Geological Survey, Columbia, MO); K. Liber (University of Wisconsin, Superior, WI); L. Maltby (University of Sheffield, Sheffield, U.K.); and G. van Aggelen (Environment and Climate Change Canada, North Vancouver, BC). Three of the principal authors of USEPA (1994a), namely, C.G. Ingersoll, G.T. Ankley, and G.A. Burton, served on Environment and Climate Change Canada's committee of scientific experts.

In addition to the members of the Environment and Climate Change Canada scientific advisory committee who contributed to the development of this biological test method, the following people reviewed the final draft of the first edition of this report and provided many useful comments: J. Black and W. McCulloch (EA Engineering Science and Technology Inc., Sparks, MD); D. Boersma (Commercial Chemicals Branch, Environment and Climate Change Canada, Ottawa, ON); R. Casey (Alberta Environmental Centre, Vegreville, AB); G. Dave (University of Göteberg, Sweden); M.P. Hamer (Zeneca Agrochemicals, Bracknell, Berkshire U.K.); M.L. Inman (Exxon Biomedical Sciences Inc., Millstone, NJ); R.A. Hoke (SAIC, Hackensack, NJ.); J.W. Lazorchak (USEPA, Cincinnati, OH); B. McGee (University Maryland, Queenstown, MD); M.H. Murdoch (EVS Consultants Ltd., North Vancouver, BC); C. Naylor (University of Sheffield, U.K.); A. Mueller (Federal Biological Research Center for Agriculture and Forestry, Kleinmachnow, Germany); G. Pagano (Instituto Nazionale Tumori-Pondazione, Naples, Italy); F. Quiniou (IFREMER, Centre de Brest, Plouzané, France); C. Roghair (Nat. Inst. Publ. Health and Environ. Prot., Bilthoven, Netherlands); G.L. Stephenson (Ecological Services for Planning Ltd., Guelph, ON); K. Taylor (Commercial Chemicals Branch, Environment and Climate Change Canada, Ottawa, ON); A. Verbeek (Chemex Labs. Alberta Inc., Edmonton, AB); L. Vigano (Instituto de Ricerca Sulle Acque, Milano, Italy); P. Winger (Univ. of Georgia, Athens, GA).

The second edition was prepared by J. Miller (Miller Environmental Sciences Inc., King City, ON), with assistance and guidance from L. Taylor (Manager, Method Development and Applications Unit) and L. Van der Vliet (Biological Assessment and Standardization Section) of Environment and Climate Change Canada, Ottawa, ON. Statistical guidance was provided by B. Zajdlik (Zajdlik & Associates, Rockwood, ON), and written and reviewed by L. Van der Vliet, both of whom are sincerely thanked. The second edition included numerous updates based on several studies that had taken place among certain Canadian toxicology testing laboratories in the past several years. The following were the laboratory personnel participated in these studies and/or were on the ad-hoc working group, and provided many useful comments for the second edition: K. Doe and P. Jackman from Environment and Climate Change Canada's Atlantic Laboratory for Environmental Testing (ALET, Moncton, NB); K. Hunter, T. Watson-Leung, J. Schroeder, and D. Poirier from Ontario Ministry of the Environment, Aquatic Toxicology Unit (OMOE, Toronto, ON); G. van Aggelen, C. Buday, and G. Schroeder from Environment and Climate Change Canada's Pacific and Yukon Laboratory for Environmental Testing (PYLET, North Vancouver, BC); and A. Bartlett and W. Norwood from the National Water Research Institute, Environment and Climate Change Canada (Burlington, ON). L. Novak, K. Holtze, and E. Jonczyk, all of Aquatox Testing & Consulting Inc. (Guelph, ON) are also acknowledged for the work they did on the refinement of the water-only test methodology for *Hyalella*. Thanks also to L. Taylor who acted as Scientific Authority and provided technical input and guidance throughout the preparation of the second edition report.

This (third) edition was prepared by Lesley Novak (AquaTox Testing & Consulting Inc., Guelph, ON), and Dr. Lisa Taylor (Manager, Method Development and Applications Unit) and includes the addition of the 42-day exposure that is based on studies that have taken place over the past several years. Grateful acknowledgement is made of the following laboratory personnel who participated in these studies (including the inter-laboratory studies) and also provided useful comments during various stages of method development: Lesley Novak, Martina Rendas and Conrad Neufeld from AquaTox Testing & Consulting Inc.(Guelph, ON), Paula Jackman from Environment and Climate

Change Canada's Atlantic Laboratory for Environmental Testing (ALET; Moncton, NB), Grant Schroeder, from Environment and Climate Change Canada's Pacific & Yukon Laboratory for Environmental Testing (PYLET; North Vancouver, BC), Marriah Grey and Janet Pickard from Maxxam Analytics (Burnaby, BC), James Elphick from Nautilus Environmental (Burnaby, BC), Trudy Watson-Leung, Kim Mahon and Lisa Kennedy from Ontario Ministry of the Environment and Climate Change (Toronto, ON), Teresa Norberg-King and David Mount from U.S. Environmental Protection Agency (Duluth, Minnesota, Christopher Ingersoll from Columbia Environmental Research Center, U.S. Geological Survey (Columbia, MO), and David John Riecks-Soucek and Amy Dickinson from Illinois Natural History Survey, University of Illinois at Urbana-Champaign (Champaign, IL). Thanks also to Leana Van der Vliet (Environment and Climate Change Canada; Ottawa, ON) who provided review comments on the test design.

Section 1

Introduction

1.1 Background

Aquatic *toxicity tests* are used within Canada and elsewhere to determine and monitor the *toxic* effects of discrete *substances* or complex mixtures that *might* be harmful to indigenous aquatic life in the environment (water and *sediment*). The results of toxicity tests *can* be used to determine the need for control of discharges, to set effluent standards, and for research and other purposes. Recognizing that no single test method or test organism can be expected to satisfy a comprehensive approach to environmental conservation and protection, the Inter-Governmental Ecotoxicological Testing Group (IGETG) (Appendix B) proposed to develop a set of aquatic toxicity tests that would be broadly acceptable for use in Canada, and would measure different toxic effects using different test *substances* or *materials* (e.g., samples of *chemical* or chemical substance, *effluent*, *receiving water*, or sediment) and organisms representing different trophic levels and taxonomic groups.

In 1987, Environment and Climate Change Canada and the IGETG recommended that a consistent set of sediment testing methods be developed for routine use in preventing, appraising, remediating, and managing *contaminated sediment* (Sergy, 1987). At that time, Environment and Climate Change Canada's regional laboratories (Appendix C) began a series of studies (McLeay *et al.*, 1989, 1991, 1992, 1993; Paine and McPherson, 1991a, b; Doe and Wade, 1992; Yee *et al.*, 1992) to develop and validate a standardized biological test method for measuring the *toxicity* of samples of contaminated sediment. The test method would use one or more of six species of marine or estuarine amphipods common to Canadian Pacific or Atlantic coastal waters (EC, 1992a). A test for survival and growth in sediment, using the freshwater amphipod *Hyalella azteca*, was one of the sediment toxicity tests selected by the IGETG members to be standardized by Environment and Climate Change Canada. In 1997, this method (EPS 1/RM/33) would become part of a series of biological test methods prepared by Environment and Climate Change Canada to help meet Canadian requirements related to environmental appraisal and protection (EC, 1997b). The widespread distribution and common occurrence of *H. azteca* in association with freshwater sediment, together with its ecological importance, ease of culturing and handling during testing, rapid growth, short life cycle, sensitivity to contaminants in sediment, and extensive use in sediment toxicity tests, led to this selection. After 9 years of application by private and public sector testing laboratories, Environment and Climate Change Canada recognized that specific aspects of the test method needed to be re-evaluated. In 2006 Environment and Climate Change Canada established an ad-hoc working group to review EPS 1/RM/33 and to determine research priorities. The goal was to develop a research plan to address some methodology issues, and to come up with recommendations to be included in the method revision. Since then several studies have taken place among certain Canadian toxicology testing laboratories with research focused on improving *Hyalella* testing parameters such as feeding rates (type of food and ration), the age of test organisms used and growth variability, light intensity for culturing, ratios of water to

sediment used in the test, and a 14-day *water-only* exposure. The results of those studies are presented in a report (MESI, 2010), and revisions based on its conclusions were included in a second edition of EPS 1/RM/33 (2013). Further revisions in the second edition included updated statistical guidance (with the exclusion of laboratory replicates [i.e., replicate test vessels] for single-concentration tests of field-collected sediments) and new options for type of exposure (i.e., *static* and triggered *static-renewal*).

In 2010, Environment and Climate Change Canada began investigating a reproduction endpoint over a longer term exposure (42 days). The entire 42-day test involves continuous exposure to sediment and eliminates the potential for depuration and loss of sensitivity, which can occur in other standardized methods (e.g., US EPA, ASTM) where adults are transferred to a water-only exposure on Day 28 before release of their first brood. The mid-test transfer to clean water was a feature based on the results of a recovery trial conducted by Tomasovic *et al*., (1995) which concluded that reproductive endpoints could be biased because of low recovery of young amphipods from sediment. However, inter-laboratory testing by Environment and Climate Change Canada has shown that with experience and use of specific recovery techniques (e.g., use of light table for a contrasting background) 88% of young amphipods (2 to 5 days old) can be recovered with low variability (*Coefficient of Variation* (CV) ~ 8.6 %) (Taylor and Novak, 2016). There are several reasons why acceptable recovery is attainable with this new 42-day test method. First, the volume of sediment is much lower (1/5th) relative to that used by Tomasovic *et. al*. (1995) (18 mL vs. 100 mL). Second, the absence of sediment sieving in this method reduces injury or death of young. Finally, food ration is increased and ramped up over time (Ivey *et. al.*, 2016), allowing adults to produce two broods that grow to a visible (without the aid of a microscope or magnifying glass) size within the 42 day exposure. Soucek *et al*. (2016) has also shown that control *H. azteca* fed improved diets can have their first brood as early as day 24, with the second brood occurring between days 33 to 35. Therefore, in the 42-day test described herein, the youngest control organisms would be approximately a week old by test completion.

Because organism recovery at the end of the 42-day test is critical to success of this continuous sediment exposure method, technicians conducting this test will be required to prove recovery proficiency prior to performing the test (also called a technician performance criterion). The recommended average recovery (i.e., recovery that labs should strive to achieve) will be 85%, with a minimum acceptable recovery of 80% (Taylor *et al.*, 2016; Taylor and Novak, 2016).

The current report represents a revised and updated version of EPS 1/RM/33 and is intended to supersede and replace guidance for testing survival and growth in sediment using the freshwater amphipod *Hyalella azteca* provided in Environment Canada's earlier versions of Report EPS 1/RM/33 (2013).

Tests with samples of freshwater sediment and the freshwater amphipod *H. azteca* have historically been carried out by Canadian investigators using various procedures including those published by Borgmann and Munawar (1989), Borgmann *et al.* (1989), ASTM (1991a, 1993), as well as the unpublished standard operating procedure of the National Water Research Institute (NWRI, 1992). Other notable procedures for culturing and testing *H. azteca*, which have influenced

the preparation of this biological test method, include: de March, 1981; FDA, 1987; Ingersoll and Nelson, 1990; Smith *et al.*, 1991a, b; USEPA, 1991a, b, 1994a, 2000; DFO, 1992; Norberg-King, 1992; Ankley *et al.*, 1993a; Brooke *et al.*, 1993; Kubitz, 1993a, b; Borgmann, 1996, 2002; Borgmann and Borgmann, 1997; Borgmann and Norwood, 1999; Ivey *et al*., 2004, 2011, 2016; Borgmann *et al.,* 2005a, b; AFNOR (2005); ISO (2011); Hockett *et al*. (2011); P. Jackman (ALET, Environment Canada, Moncton, NB, personal communication, 2012).

In 1994, the United States Environmental Protection Agency (USEPA) published new methods for measuring the toxicity of sediment associated contaminants that include a *solid-phase sediment* toxicity test using *H. azteca*[1] (USEPA, 1994a). These sediment assays were updated in 2000 (USEPA, 2000) and have been adopted as standard test methods by Committee E47 of the American Society for Testing and Materials (ASTM, 1995a, 2010). This biological test method developed by Environment and Climate Change Canada relies heavily on the specific procedures for culturing and testing *H. azteca* detailed in USEPA (1994a, 2000). A significant distinction, though, is that the toxicity test using *H. azteca* described in USEPA (1994a, 2000) and ASTM (1995a, 2010) is a 10-day assay for effects on amphipod survival, whereas Environment and Climate Change Canada's method includes a 14-day *survival-and-growth* test. Additionally, the 14-day test is normally carried out as a *static* exposure (rather than the daily-renewal exposure described in USEPA 1994a and 2000) in which the overlying water is not renewed during the test (except for replacing losses due to evaporation), and is aerated continuously. A *static-renewal* exposure for the 14-day survival and growth test has also been described herein (see Sections 3.1 and 4) and can be triggered by the fouling of *test water* overlying reference sediment (i.e., due to high levels of *ammonia*, *pH* levels outside the tolerance range of *Hyalella azteca*, and/or low levels of dissolved oxygen). The static-renewal exposure during the 14-day test can be used if the objective of the test is to assess toxic effects due to substances or materials without the deleterious or modifying effect of these confounding factors.

The new methodology for the 42-day exposure described herein also eliminates the potential for depuration and loss of sensitivity by conducting the entire exposure in sediment; in contrast to existing standardized methods where adults are transferred to a water-only exposure before release of their first brood at Day 28. This mid-transfer to clean water was due to the results of a juvenile *H. azteca* recovery trial conducted in the 1990s (Tomasovic et al., 1995), which concluded that reproductive endpoints could be biased because of low recovery of young amphipods from sediment. However, in this Environment and Climate Change Canada test method, key to successfully conducting the longer-term test is the recovery from sediment and accurate counting of the young of the adults at the end of the 42 day exposure (Taylor and Novak, 2016).

The first edition of this method included an option for a daily-renewal test in keeping with USEPA (1994a) and ASTM (1995a), which

[1] USEPA (1994a) was prepared by members of the United States Freshwater Sediment Toxicity Assessment Committee, and reflects a consensus opinion of U.S. and Canadian researchers actively engaged in sediment toxicity tests using *H. azteca* (Ingersoll, 1992; Norberg-King, 1992; Burton and Ingersoll, 1994; Ingersoll *et al.*, 1995; Burton *et al.*, 1996). The USEPA published a revised version (i.e., second edition) of the manual in March 2000 (USEPA, 2000).

required twice-daily renewal of the *overlying water* in test vessels, and normally no aeration of the overlying water. This option is still included herein for the 14-day test, but for use only under specific conditions, described in Section 4.

When conducting the 42-d survival, growth and reproduction test (Section 8), the overlying water is continuously aerated and renewed on days 14, 28 and 35. This frequency of water renewal was based on method development studies which showed water quality deterioration was only observed at Day 35.

For the 14-d exposures, the animals are fed either a standardized mixture of yeast, cereal grass media, and trout chow (YCT), a commercial flaked fish food, or a 1:1 mixture of YCT and flaked fish food, either daily, or three times per week on non-consecutive days. For the 14-d exposures, the ration provided is kept constant throughout the exposure. In comparison, during the 42-d exposure, the animals are fed YCT and flaked fish food three times per week, but with a gradual increasing ration to promote growth and subsequent reproduction.

Universal procedures for preparing and conducting sediment toxicity tests using *H. azteca* are described in this revised and updated method. Also presented are specific conditions and procedures that are required or recommended when using the test for evaluating different types of substances or materials (e.g., samples of field-collected sediment or particulate waste, or samples of one or more chemicals or chemical substances experimentally mixed into or placed in contact with natural or formulated sediment). This updated version of EPS 1/RM/33 also includes a 14-day *water-only* exposure (Section 7), which can be used either alone or with the *Hyalella* sediment test to help differentiate between historical contamination in the receiving environment and contribution of the current industry's effluent. This water-only test can be applied in "Investigation of Cause" (IOC) studies currently required under Environment and Climate Change Canada's Environmental Effects Monitoring (EEM) program.

The flowchart in Figure 1 gives a general picture of the universal topics covered herein, and lists topics specific to testing samples of field-collected sediment, similar particulate waste (e.g., sludge, drilling mud, or dredged material); sediment spiked experimentally with chemical(s), *contaminated sediment*, or particulate waste; or samples of *wastewater* or chemicals in a water-only exposure.

The first edition of this biological test method was developed following a review of variations in specific culturing and test procedures indicated in existing Canadian and

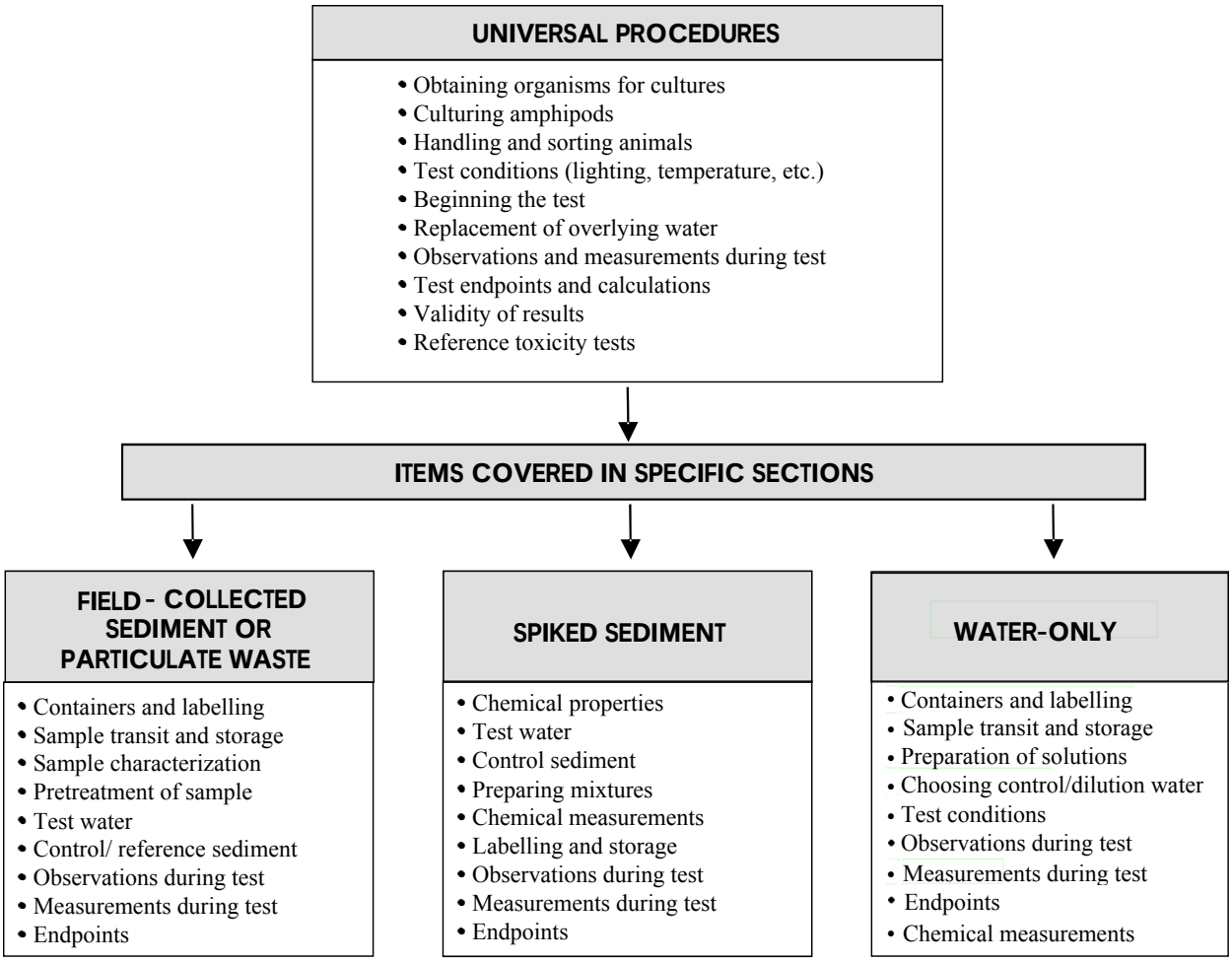


Figure 1 Considerations for Preparing and Performing Toxicity Tests Using *Hyalella azteca* and Various Types of Test Materials or Substances

United States methodology documents[2] that describe how to prepare and conduct sediment toxicity tests using the freshwater amphipod *H. azteca*. A summary of existing or past procedural variations for culturing this species and for harvesting young for use in toxicity tests is found in Appendix D. A summary of variations in existing or past procedures for conducting growth and/or survival tests for sediment toxicity using *H. azteca* is found in Appendix E. Appendix F provides a summary of interlaboratory variations in conditions and procedures for undertaking *reference toxicity tests* with *H. azteca.*

The biological *endpoints* for the 14-day survival and growth test methods described

[2] Documents used in preparing listings of procedural specific variations (see Appendices D, E, and F) include published "how-to" references, unpublished Standard Operating Procedures of governmental testing facilities, and draft reports. Citations of source documents are listed in these appendices by originating agency, rather than by author(s), although the authors and formal citations are identified in the appendices. Appendices D, E, and F have not been updated in this third edition test method document.

herein are mean percent survival and mean dry weight (as an indication of growth) at the end of the 14-day test.[3] The test methods are intended for use in evaluating the toxicity of samples of:

i) field-collected freshwater sediment (Section 5);

ii) industrial or municipal sludge and similar particulate wastes that might affect the freshwater environment (Section 5);

iii) mixtures of one or more chemicals or chemical substances within or overlying freshwater sediment (Section 6); and

iv) effluent, elutriate, leachate, receiving water, or chemicals in water-only tests (Section 7).

The specific procedures for conducting the 42-day survival, growth and reproduction test are described in Section 8.

Biological endpoints for the 42-day exposure, include the number of surviving adults (male and female), adult dry weight, number of progeny, and number of young per surviving female. Additional endpoints that could be measured include bioaccumulation (Taylor *et al.,* 2016), *biomass* and *survival-normalized reproduction* (with the latter being useful for those sediments that have both a survival and reproduction effect).

A 10-day test method for toxic effects of these materials on survival and growth of larval freshwater midges (*C. tentans* or *C. riparius*) was also developed by Environment and Climate Change Canada (Environment Canada, 1997a), and *may* be used in conjunction with or as an alternative to this test. The present sediment test method may also be used to measure and appraise the toxicity of contaminants in estuarine sediment or of chemical/sediment mixtures where the *salinity* of the overlying and/or *pore water* does not exceed 15‰ (Nebeker and Miller, 1988; USEPA, 1994a, 2000). Environment and Climate Change Canada's biological test method using one or more recommended species of estuarine or marine amphipods common to Canada's Atlantic or Pacific coastal waters (EC, 1992a) is normally used for measuring and evaluating the toxicity of contaminants in estuarine or marine sediment.

In formulating these procedures, an attempt has been made to balance scientific, practical, and cost considerations, and to ensure that the results will be sufficiently precise for most situations in which they will be applied. It is assumed that the user has a certain degree of familiarity with aquatic toxicity tests. Explicit instructions that might be required in a regulatory *protocol* are not provided in this report, although it is intended as a guidance document useful for that and other applications.

[3] It has been common practice to describe biological endpoints that measure weight as "growth." However, if the organism weights at the start of the exposure period (initial weights) have not been subtracted from the weights at the end of the exposure period (final weights), then "growth" is not an accurate term. Because there is no correction for initial weights with this *Hyalella* test, the correct measurement endpoint is "final weight." After an evaluation of sample data sets and possible impacts, the Method Development and Applications Unit decided to continue current practice (i.e., not correct for initial dry weights). In this revised edition, the word "growth" has been replaced with "(final) dry weight," where practical; however, in the context of this document, both "terms" refer to the mean dry weight of *Hyalella* at the end of the test. There has been no change in the measurement or calculation of the biological endpoint; the terminology has simply been revised to accurately describe the measurements.

For guidance on the implementation of this and other biological test methods, and on the interpretation and application of the endpoint data, the reader *should* consult Environment Canada (1999a).

1.2 Identification, Distribution, and Life History

H. azteca (Saussure) is an epibenthic, detritus-feeding, sediment-burrowing, freshwater amphipod. The distinguishing features of this small (male to 8 mm, female to 6 mm) crustacean species are described and illustrated in Appendix G.[4]

H. azteca resides in temperate lakes, ponds, and slow-flowing streams, in close association with the surficial 1 or 2 cm of sediment. The species has achieved densities of >10 000/m^2 in preferred habitats and can also be found in lower numbers in sloughs, ditches, springs, rivers, and marshes (USEPA, 1994a, 2000). The species is widely distributed on the North American continent, and has been found in surficial sediments of lakes from Guatemala and the Caribbean Islands to Inuvik, Northwest Territories (de March, 1981). Information on the known distribution of *H. azteca* in Canadian waters, and its habitat, is included in Appendix G.

Amphipods are an abundant component of benthic communities in North American freshwater, estuarine, and marine environments. Freshwater amphipods including *H. azteca* are an important source of food for many species of fish, waterfowl, wading birds, salamanders, and larger invertebrates (de March, 1981). *H. azteca* is reported to selectively ingest bacteria and algae that adhere to sediment particles <65 µm (Hargrave, 1970). The animal reproduces sexually, and the eggs and live young are carried in a brood pouch on the female's ventral surface. Immediately before mating, the female molts and releases its offspring from the previous mating. Depending on the size and condition of the female, 1 to 50 offspring can be produced and released at each molt. At 25°C, *H. azteca* reaches sexual maturity in about 28 to 33 days (de March, 1981; USEPA, 1991a, b). Studies by Soucek *et al*. (2016) have shown that that control *H azteca* fed improved diets have their first brood as early as day 24 and the second brood occurring between days 33-35. For further details on the life history of this species, see Appendix G.

1.3 Historical Use in Toxicity Tests

Burton (1991) provides an excellent review of the various methods used historically for measuring the toxicity of freshwater sediments, including toxicity tests with *H. azteca* and other species of freshwater amphipods. Nebeker *et al.* (1984) first recommended that *H. azteca* be used in partial life-cycle tests to measure the toxicity of contaminated freshwater sediment. Nebeker and Miller (1988) demonstrated that this species of amphipod will survive and reproduce in toxicity tests with sediments from estuarine sites when the sediment in the test vessels is overlain by fresh water. Subsequent studies have demonstrated that *H. azteca* can be *cultured* in water with a salinity of up to 15‰, then used in toxicity tests associated with estuarine discharges (Ingersoll *et al.*, 1992) or contaminated estuarine sediments (McGee *et al*., 1993). The USEPA and the United States Army Corps of Engineers (USACE) have recommended that *H. azteca* be used for

[4] Evidence from molecular analysis (e.g., DNA barcoding) has suggested that there are likely to be numerous species within *Hyalella azteca* (Witt and Hebert, 2000), even though morphologically these species are identical to each other (see Section 2.1).

evaluating dredged material proposed for discharge in inland and coastal waters (USEPA/USACE, 1994). Ingersoll *et al.* (1995) reviewed methods and applications for sediment toxicity tests using *H. azteca*.

H. azteca has been used in *water-only* acute and chronic toxicity tests with various chemicals (FDA, 1987; Borgmann and Munawar, 1989; Borgmann *et al.*, 1989, 1990, 1991, 1993, 2005b, c, 2007; Schubauer-Berigan *et al.*, 1993; Borgmann, 1994; Hoke *et al.*, 1995; Phipps *et al.*, 1995; Borgmann and Borgmann 1997; Wang *et al*., 2008). Biological endpoints for these tests have included survival, growth, and reproductive success in partial or full (12 to 14 weeks at 25°C) life-cycle studies. Water-only tests for bioaccumulation of specific chemicals have also been conducted using *H. azteca* (Borgmann *et al.*, 1990, 1991, 1993, 2010; Norwood *et al*., 2007a, b). The toxicity and bioaccumulation of chemicals added to sediment (*spiked sediment* tests) have been studied by several researchers using *H. azteca* (Landrum and Scavia, 1983; Cairns *et al.*, 1984; Nebeker *et al.*, 1986, 1989; Smith *et al.*, 1992a; Suedel *et al.*, 1993a, b; Kubitz *et al.*, 1995; Milani *et al.*, 1996; Whiteman *et al.*, 1996; Besser *et al*., 1998; Borgmann *et al.,* 2001a, b; Bartlett *et al*., 2004, 2007; Nowierski *et al*., 2005; Norwood *et al*., 2009).

Many investigators have successfully used *H azteca* to appraise the toxicity of whole (*solid phase*) samples of freshwater sediment (e.g., Nebeker *et al.*, 1984; Borgmann and Munawar, 1989; Burton *et al.*, 1989; Ingersoll and Nelson, 1990; Ankley *et al.*, 1993a, b; Borgmann and Norwood, 1993; Kubitz, 1993a; Kubitz *et al.*, 1993; Sibley *et al.*, 1993; West *et al.*, 1993; Burton and Ingersoll, 1994; Kemble *et al.*, 1994; Pastorok *et al.*, 1994; Becker *et al.*, 1995; Ingersoll *et al.*, 1995; Kubitz *et al.*, 1995, 1996; Reynoldson *et al.*, 1995; Burton *et al.*, 1996; Milani *et al.*, 1996; Borgmann *et al.,* 2001a, 2004; Borgmann and Norwood, 2002; Bartlett *et al.,* 2005). Biological endpoints for these tests, which can be used to evaluate the spatial and temporal variability in toxicity of samples of field-collected sediment, are typically mean percent survival and mean growth (length or weight) at the end of the test. Toxicity tests using aqueous extracts (pore water and/or elutriates) of freshwater sediments have also been conducted using *H. azteca* (Burton *et al.*, 1989; Ankley *et al.*, 1991; Schubauer-Berigan and Ankley, 1991; Sibley *et al.*, 1993).

Results of laboratory sediment toxicity tests using *H. azteca* have been examined to assess their worth in identifying sites where natural populations of benthic organisms are affected by toxic sediment contaminants (Becker and Bigham, 1993; Burton and Ingersoll, 1994; Canfield *et al.*, 1994; Schlekat *et al.*, 1994). Such *field validation* studies typically integrate the findings of the laboratory toxicity tests with concurrent chemical analyses of sediment samples and field surveys of the diversity and abundance of benthic communities, using a Sediment Quality Triad (Chapman *et al.*, 1986, 1987, 1991) or similar approach. To date, integrated laboratory and field studies that include sediment toxicity tests using *H. azteca* or other species of amphipods have indicated that these tests can provide reliable evidence of biologically adverse contamination of sediment in the field (Swartz *et al.*, 1982, 1985a, 1986, 1994; Becker *et al*., 1990; Canfield *et al.*, 1994; USEPA, 1994a; Day *et al.*, 1995a; Borgmann *et al*., 2001b, 2004).

At the time of the release of this third edition, the US EPA was in the process of revising and updating their 42-day test, to include recent findings from various research

studies including Ivey *et al.* (2016) and Soucek *et al*. (2016).

1.4 Laboratory Tolerance and Relative Sensitivity

A number of studies have examined the tolerance of *H. azteca* to certain natural environmental variables under laboratory conditions. Effects of temperature on the tolerance, behaviour, and reproductive biology of this animal are summarized in de March (1981). The natural range of temperatures tolerated by *H. azteca* is 0 to 33°C. Generally, maximum numbers of young are produced between 26 and 28°C, whereas temperatures of 33 to 37°C are *lethal*. Temperatures of 0 to 10°C cause complete immobility; temperatures of 10 to 18°C delay maturation, result in a low rate of reproduction, and produce large adults; and temperatures ≥20°C decrease maturation time, increase the rate of reproduction, and produce small adults (de March, 1981).

Reproduction of *H. azteca* in the laboratory is successful and continuous if a *photoperiod* of at least 16 hours of light per day (≥16L:8D) is used. Shorter daylight hours (≤12L:12D) can result in a reproductive resting stage (de March, 1977). *H. azteca* has been cultured successfully using broad-spectrum fluorescent lighting with an intensity of about 500 to 1000 *lux* (Ingersoll and Nelson, 1990; Ankley *et al.*, 1991; USEPA, 1994a, 2000). Covering culture jars with aluminium foil did not affect survival, but reduced growth rates and eliminated reproduction (Borgmann *et al.*, 1989).

H. azteca can survive exposure to low levels of dissolved oxygen for extended periods. de March (1981) summarized studies indicating that this species could survive in stagnant water, and cited two independent investigations which reported a 48-h *LC50* of 0.7 mg O_2/L for this species. Nebeker *et al.* (1992) also found that this animal could survive acute or prolonged exposure to low dissolved oxygen levels; both 96-h and 30-d LC50s were <0.3 mg O_2/L. However, growth and reproduction (mean number of young) were both reduced after 30 days' exposure to water with ≤1.2 mg O_2/L. The effects of higher (but below saturation values) concentrations of dissolved oxygen on growth and reproduction of *H. azteca* were not investigated by Nebeker *et al.* (1992).

Little information is available on how pH affects the survival of *H. azteca.* de March (1979) reported that survival of this species was optimum at pH values of 6 to 8, and that pH values ranging from 4 to 5 resulted in gradual mortalities. It is not known if there is any information on the effects of fresh water with differing pH values on the growth rates or reproductive success of this amphipod.

There is little definitive information on the influence of water hardness or alkalinity on the well-being of *H. azteca.* The USEPA (1991b) observed that the reproductive success of this species was often poor when cultured in *reconstituted water* adjusted to low hardness values using conventional recipes (e.g., those in USEPA, 1985a, b, 1991c). However, this problem might have been due to a chemical imbalance of the ions in solution for this species, rather than adverse hardness *per se*; and a recipe for preparing reconstituted water with a hardness of 90 to 100 mg $CaCO_3$/L yielded better (although not universal) success and was recommended in the first edition of this test method document (USEPA, 1994a). For longer tests and culturing, however, it provided poor and inconsistent results (Borgmann, 2002). Since then, a five-salt standard artificial medium (SAM-5S), developed at the NWRI (Borgmann, 1996)

has had more universal success (Borgmann and Borgmann, 1997; Borgmann, 2002; Ivey *et al*., 2004; Borgmann *et al*., 2005b, 2010) and is now recommended in this third edition test method document. However, the recommended bromide concentration has been reduced (Br concentration of 0.02 mg/L, added as NaBr), and now also includes a minimum chloride concentration (chloride concentration of 15 mg/L, added as NaCl) (Ivey *et al.*, 2016). The SAM-5S *reconstituted water* contains bromine (Br), which in a specific ratio with calcium (Ca) has been found to be required for effective utilization of calcium by *Hyalella*. The medium also contains Na and HCO_3, which are the most essential ions for *H. azteca* survival, and Mg and K, which are needed for optimal growth and reproduction (Borgmann, 1996; See Section 2.3.4). Further research is required to determine the ranges of water hardness and alkalinity that are suitable for culturing and testing *H. azteca*.

H. azteca has been shown to be euryhaline, and the species has been successfully cultured and/or tested using estuarine water and estuarine sediment. In tests using organisms acclimated to fresh water, Nebeker and Miller (1988) reported that, depending on age of the test organisms (i.e., young adults or mature adults), 10-d LC50s for *H. azteca* exposed to various salinity concentrations ranged from 19 to 24‰ and 24-h LC50s for mature adults ranged from 16 to 19‰. For organisms acclimated to fresh water, inhibition of reproductive success (i.e., number of young produced) was evident at salinities of 10.4‰ and higher (Nebeker and Miller, 1988). Presumably, the salinity tolerance of this species might be greater if the animals are acclimated to estuarine water before testing. de March (1981) noted that *H. azteca* can survive a salinity as high as 30‰ if acclimated gradually. Other studies have demonstrated that *H. azteca* can be cultured successfully in water with a salinity of 15‰ or less (McGee *et al.*, 1993; USEPA, 1994a, 2000). More recent research shows that not all strains of *Hyalella* have the same tolerance to salinity, and therefore when conducting tests at higher salinities, the strain of *Hyalella* needs to be selected carefully (Borgmann, 2002).

The influence of natural physicochemical properties of sediments on the performance of *H. azteca* in sediment toxicity tests has been examined. Ingersoll and Nelson (1990) found that this species has an extremely wide tolerance of sediment grain size. In long-term exposures to *clean sediments* ranging from >90% silt- and clay-sized particles to 100% sand-sized particles, no detrimental effects on either survival or growth were noted. Similarly, Ankley *et al.* (1994) conducted 10-day *H. azteca* sediment assays using 50 uncontaminated samples of lake sediment with particle sizes ranging from 95% clay to 100% sand, and organic carbon content from 0.3 to 8.1%. These researchers found no correlation between amphipod survival rates and sediment characteristics including particle size, organic carbon content, or mineralogical composition, provided the animals were fed during the tests. In 10-day survival tests with laboratory-formulated or clean field-collected sediments, Suedel and Rodgers (1994a) determined that *H. azteca* was tolerant of all of the sediment particle size distributions (0 to 100% sand, 0 to 100% silt, and 0 to 60% clay) and ranges of organic carbon content (0.1 to 8.0%) examined. In 48-h tests with sediment spiked using a range of concentrations of alkylbenzene sulphonate, Cano *et al.* (1996) found that enriching the sediment with peat moss increased the acute lethal tolerance of *H. azteca* to this surfactant when total organic carbon content was ≥1.5%.

The sensitivity of *H. azteca* to sediments or chemicals, relative to that of other freshwater species commonly used in toxicity tests, has been evaluated in a number of studies. Reviews of comparative toxicity data indicate that *H. azteca* is one of the most sensitive freshwater species (Burton, 1991; USEPA, 1994a, 2000). Acute lethality *water-only* tests with a number of industrial effluents indicated that the sensitivity of *H. azteca* was similar to that of rainbow trout, *Oncorhynchus mykiss* (Maciorowski, 1975). Similarly, results for comparative 96-h (*H. azteca*) or 48-h (*Ceriodaphnia dubia*) water-only tests using potassium chloride (KCl) showed that the acute lethal tolerance of *H. azteca* and the freshwater daphnid *C. dubia* to this *reference toxicant* was similar (Smith *et al.*, 1991b). In comparative acute lethality tests with sediment pore water or elutriate, *H. azteca* was either as, or slightly more, sensitive than *C. dubia* or larval fathead minnows (*Pimephales promelas*), with the oligochaete *Lumbriculus variegatus* being the least sensitive (Ankley *et al.,* 1991). Chronic water-only tests with cadmium and pentachlorophenol, using *H. azteca*, another amphipod (*Gammarus fasciatus*), or *Daphnia magna* as test organisms, showed that the sensitivities of the two amphipod species were similar and that each species was as or more sensitive than *D. magna* (Borgmann *et al.*, 1989). Comparative 10-day water-only and *spiked sediment* tests with fluoranthene, using *H. azteca*, larval freshwater midges (*C tentans*), and *D. magna*, showed that *H. azteca* and *C. tentans* were twice as sensitive as *D. magna* in the water-only tests, whereas *H. azteca* was as or more sensitive than the other two species in the spiked sediment tests (Suedel *et al*., 1993a). In comparative water-only LC50s involving *H. azteca, C. tentans,* and *L. variegatus* exposed separately to each of five metals and five pesticides, Phipps *et al.* (1995) found that no one species was consistently most sensitive to all *toxicants*, although *H. azteca* was the species most sensitive to all five metals. Additionally, these investigators compared the LC50s derived for *H. azteca* to published toxicity values for other aquatic species exposed to the same chemicals, and noted that *H. azteca* was frequently amongst the most sensitive species (Phipps *et al.*, 1995). Comparative tests by Kubitz *et al.* (1995), using copper-spiked sediment or samples of field-collected sediment, 48-h pore water tests for survival of *D. magna* or *C. dubia*, 1-h tests for enzyme inhibition using *D. magna*, and 14-day sediment *survival-and-growth* tests with *H. azteca*, indicated that both the growth (i.e., dry weight) endpoint using *H. azteca* and the enzyme inhibition endpoint using *D. magna* were more sensitive than any of the survival endpoints.

In comparative 10-day *whole sediment* tests with field-collected sediment, West *et al.* (1993) found that, of the species compared (i.e., *H. azteca*, *C. tentans*, and *L. variegatus*), *H. azteca* was the most sensitive. As cited in USEPA (1994a, 2000), Kemble *et al.* (1994) compared the sensitivity of *H. azteca*, *C. riparius, D. magna*, and rainbow trout to samples of metal-contaminated sediment. Using length, sexual maturation, and survival as endpoints in 28-day tests, results showed that *H. azteca* was the most sensitive of the four species tested (and that length was the most sensitive endpoint). In a separate study of contaminated sediment from the Great Lakes, *H. azteca* was amongst the most sensitive and discriminatory of 24 species tested (Burton and Ingersoll, 1994; USEPA, 1994a, 2000). The results of studies by Smith *et al.* (1993) indicate that larval fathead minnows (7-day test) might be more sensitive than *H. azteca* to certain natural sediments contaminated with metals and metalloids (selenium).

Section 2

Test Organisms

2.1 Species and Life Stage

The freshwater amphipod *Hyalella azteca* ***must*** be used in this biological test method. This crustacean species is an epibenthic, sediment-burrowing detritivore that lives in close contact with freshwater sediments. Confirmation and documentation[5] of the species of test organisms received from a supplier must be made by a qualified taxonomist, at least once for any shipments of *Hyalella* provided by that supplier, using the distinguishing taxonomic features described and illustrated in taxonomic keys and in Appendix G, or using DNA-based taxonomic identification (i.e., barcoding).[6] Thereafter, periodic confirmation of the species can be made by the testing laboratory by comparing an organism from a given *batch* to a representative specimen previously confirmed to species by a taxonomist, and maintained as a preserved specimen at that laboratory (EC, 1999b), or by submitting samples of test organisms for DNA barcoding.

Juvenile *H. azteca* that are 2- to 9-days old must be used for the 14-day survival and growth test method. For greater standardization and reduced variability of final dry weight measurements, test organisms must not vary in age by more than 3 days, however an even closer age-range (i.e., ≤2 days) is highly recommended (see Section 2.3.10).[7]

For the 42-day survival, growth and reproduction test, organisms must be ≥7- to ≤9-days old. Starting mean dry weight of the 7 to 9 day old amphipods should be 0.020 – 0.035 mg/individual and should be determined at test initiation. The information on initial weights would provide added confidence that control organisms will achieve adequate growth in the control sediment to reproduce and meet the test validity criteria of ≥6.0 young per surviving female by test completion (see Section 8.5). Method development work has shown reduced reproduction in 42 days when initial mean dry weights are less than 0.02 mg/individual. Measurement of initial weights

[5] Acceptable forms of documentation include: certification from the test organism supplier, identification of laboratory specimens by a qualified taxonomist, and identification of laboratory specimens by molecular analysis (such as DNA barcoding).

[6] Using standard taxonomic keys, all *Hyalella* in Canada are expected to be identified as *Hyalella azteca* (Borgmann, 2002). However, more recent evidence from molecular analysis has suggested that there are likely to be numerous species within *Hyalella azteca* (Witt and Hebert, 2000), which on a morphological basis would be virtually identical to each other. As an interim measure, species taxonomy for *Hyalella* must be confirmed microscopically to the species level as *Hyalella azteca*, or using molecular techniques, must be confirmed to a species of *Hyalella* known to be present in North America that is closely related to *Hyalella azteca* and excludes other well-described *Hyalella* spp. (for example the use of *Hyalella montezuma*, a planktonic filter feeder, would not be permitted). Taxonomic guidance will be updated as the molecular evidence continues to coalesce. In Canadian laboratories, species have been identified as part of the "US Laboratory Clade" or the rare "Burlington Clade" (Major *et al.,* 2013). Sensitivity differences between these two clades have been reported for the metals copper and nickel (Leung *et al*., 2016).

[7] The USEPA (2000) recommends starting a *Hyalella* test with juveniles that have a narrow range in size or age (i.e., 1- to 2-day range in age) in order to reduce the potential variability in growth at the end of the test. This narrow age-range (i.e., ≤2days) is a requirement in the ISO *Hyalella* Standard 16303 (ISO, 2013).

is strongly recommended for labs first using this 42-d exposure.

2.2 Source and Acclimation

All amphipods used in a test must be derived from the same population. Sources of animals required to establish *cultures* (see Section 2.3) may be government or private laboratories, which are culturing *H. azteca* for sediment toxicity tests, or a commercial biological supplier.[8] Breeding stock can be acquired from the following Canadian sources:

Atlantic Laboratory for Environmental Testing (ALET)
Atlantic Environmental Science Centre
Environment and Climate Change Canada
P.O. Box 23005
Moncton NB E1A 6S8

Pacific and Yukon Laboratory for Environmental Testing (PYLET)
Pacific Environmental Science Centre
Environment and Climate Change Canada
2645 Dollarton Highway
North Vancouver BC V7H 1B1
Persons searching United States sources for test organisms should refer to USEPA (2000; Table 10.1) for contacts.
Breeding stock should be transported to the laboratory using the source of water in which the organisms have been reared. Water used for transporting animals should be well oxygenated (90 to 100% saturated) before shipment, and suitable substrate should be provided (Section 2.3). Shipping containers should be insulated to minimize changes in water temperature during transit. Live organisms should be transported as rapidly as is possible/practical (i.e., within 24 h). Excessive crowding of animals during shipment should be avoided to minimize stress and prevent oxygen deficiency in transit.

Upon arrival at the laboratory, organisms may be held in the water used in transit while temperature adjustments are made, or they may be transferred to well-oxygenated culture water adjusted to the temperature of the water in the shipping container. Gradual exposure of organisms to culture water is recommended in instances where there is a marked difference in quality (e.g., hardness, pH, *conductivity*) from that to which they were previously acclimated. Guidance given in Section 4.1 for acclimating organisms to test water might also be followed here when transferring amphipods from another source to culture water.

Water temperature should be adjusted gradually to the temperature specified for culturing (23°C; Section 2.3), at a rate not exceeding 2°C/day (USEPA, 1994a, 2000). During this *acclimation* period, water used to hold the breeding stock should be aerated gently. Other conditions during this interim holding period for acclimation of breeding stock to laboratory conditions should be as similar as possible to those used for maintaining cultures (Section 2.3).

It is strongly recommended that the test organisms be obtained from an in-house *culture* (see Section 2.3). If necessary, however, the test organisms can be imported (as juvenile amphipods if the organisms are 2- to 9-days old at the start of the test, and

[8] Investigators might be concerned with the effects of excessive inbreeding of laboratory cultures, or might wish to use progeny reproduced from organisms that occupied a particular locale. Accordingly, cultures may also be established using wild populations. If animals are obtained from a wild population, their taxonomy should be confirmed, and they should be cultured through several generations and evaluated for sensitivity to reference toxicant(s) before the progeny are used in toxicity tests. Obtaining wild populations of organisms for testing should be avoided unless the ability of the wild population to cross-breed with existing laboratory populations has been demonstrated (USEPA, 1994a, 2000).

ranging in age by ≤3 days) for use in the 14-day survival and growth tests, provided that Environment and Climate Change Canada's recommended procedures for the importation of test organisms for sublethal toxicity testing (EC, 1999b) are consulted and the guidance therein followed. In particular, the required conditions and procedures for the importation of test organisms, described therein, and in the following paragraphs must be adhered to, and where applicable, the requirements for in-house cultures, described herein, must also be met. In this case, each shipment or group imported would represent a discrete *batch* of test organisms. In addition, if imported for use in the 42-d survival, growth and reproduction test method, organisms (7- to 9-days old at test initiation and therefore 6- to 8-days old on arrival at the lab) must have a minimum acclimation time of 24 hours prior to use, and less than 20% mortality during the 24 hour period immediately preceding the test.[9]

If it is necessary to import test organisms, it is recommended that they be transported as young as possible to enable sufficient time for acclimation of the juveniles to laboratory test water at the test temperature (i.e., 23 ± 1°C; see Section 4), before use in the test. Each shipment of imported test organisms must include a written statement that identifies the number and source, as well as the age of the juveniles shipped, and the date and time of that shipment. The organisms must be from a dedicated culture (i.e., supplier that maintains ongoing cultures) that have met the health criteria and quality assurance requirements outlined herein. If test organisms are imported to a testing laboratory, they must be in good health, and the mortality rate for juvenile *Hyalella* must not exceed 20% in the 24-h period immediately preceding the test (EC, 1999b). Confirmation that this mortality rate is not exceeded requires a count of the total number of *Hyalella* (live and dead) received from the supplier and a count of surviving *Hyalella* in the 24-h period just prior to their transfer to test vessels. Any requirement for *monitoring* water quality characteristics and other culture conditions (e.g., temperature) must be followed by the supplier, as specified in this test method document. The testing laboratory must establish an in-house system for evaluating the health of each shipment of organisms.

In each instance where juvenile test organisms are imported to the testing laboratory, the temperature and dissolved oxygen concentration in the water within the shipping container(s) must be measured and recorded upon departure from the supplier's facility, as well as upon arrival at the testing laboratory (EC, 1999b). During transportation, the temperature of this water should be maintained at or near the required test conditions, and should not change by more than 3°C during transit.

Additionally, the dissolved oxygen content must be ≥80% saturation upon (EC, 1999b). Water used for transporting test organisms must be well oxygenated (e.g., 90 to 100% saturation) before shipment.

Upon arrival at the testing laboratory, the organisms must be acclimated as gradually as possible to the laboratory holding and/or testing conditions such that the organisms are not stressed. Holding conditions must be the same as the test conditions with respect to critical factors such as temperature, light and *photoperiod*. Test organisms that will be used within the first 24 to 48 hours after arrival at

[9] Inter-laboratory testing showed the use of imported organisms introduced some variability in recovery results, but acclimation of at least 24-hours appeared to improve the health of young amphipods (Taylor and Novak, 2016).

the testing facility should be cultured by the supplier in water that has similar qualities (temperature, pH alkalinity, hardness, etc.) as the laboratory's water (i.e., that to be used as *overlying water* in the test). Gradual exposure of organisms to the testing-laboratory's water is recommended in all cases, but especially in instances where there is a marked difference in quality from that to which they were previously acclimated. Guidance provided earlier in this section and in Section 4.1, should be followed when transferring *Hyalella* from one water source to another. Acclimation should be started upon arrival at the testing facility, and should be completed 2 days prior to setting up a test (EC, 1999b).

2.3 Culturing

2.3.1 General

General guidance and recommendations for culturing *H. azteca* in preparation for sediment and *water-only* toxicity tests are provided here. In keeping with the premise "What might work well for one laboratory might not work as well for another laboratory" (USEPA, 1994a, 2000), explicit directions regarding many aspects of culturing, including the choice of water-renewal conditions, substrate for amphipods, and food type and ration, are left to the discretion and experience of laboratory personnel. Performance-based criteria[10] are used to evaluate the suitability of the cultured organisms for tests, and the acceptability of the test results. To be suitable for use in tests, cultures must have low mortalities, and the cultured organisms must appear healthy, exhibit normal feeding and other behaviour, be of the required age for the specified test; i.e., for the 14-day survival and growth tests, organisms must be between 2- and 9-days old, and range in age by ≤3 days (≤2 days is recommended) when a test is started; for the 42-d survival, growth and reproduction test, organisms must be between 7- and 9-days old. The acceptability of the culture should also be demonstrated by concurrent or ongoing tests using one or more *reference toxicants* (Section 4.8) and a laboratory control. If a *batch* of organisms fails to meet these criteria, the data should be carefully reviewed and depending on findings, possibly discarded.

It is the responsibility of the laboratory to demonstrate its ability to obtain consistent, precise results using one or more reference toxicants, when initially setting up to perform sediment and/or water-only toxicity tests with *H. azteca*. For this purpose, intralaboratory *precision*, expressed as a CV for the respective LC50 data (see EC, 2005 for recommended CVs for reference toxicant results), should be determined by performing 5 or more tests with different batches of test organisms from the same source, using the same reference toxicant and identical procedures and conditions for each test (see Section 4.8) (USEPA, 1994a, 2000). The laboratory should also confirm its test precision at this time by conducting 5 or more 14-day *survival-and-growth* tests using *control sediment* (for sediment tests) or *control/dilution water* (for water-only tests) and different batches of test organisms (USEPA, 1994a, 2000).[11] The conditions

[10] Performance-based criteria include those related to the survival and condition of cultured animals intended for use in the test (Section 2.3.11) as well as the criteria that must be met by control organisms for a test to be valid (Sections 4.2 and 8.14), and those related to the performance of groups of animals in reference toxicity tests (Section 4.8).

[11] The ongoing monitoring of *Hyalella* survival and growth in control sediment or control/dilution water can provide valuable information on the performance of test organisms, the quality of the control sediment, and the acceptability of the test conditions over time. Data (i.e., survival and growth) from control treatments used in definitive tests in the laboratory are not only used for the calculation of test results (see Sections 4.7, 5.6, 6.5, and 7.7) and to demonstrate test validity (see

and procedures used to perform these initial tests with control sediment and/or control/dilution water should be identical and according to Sections 4 and/or 7, respectively.

When routinely performing toxicity tests with *H. azteca, reference toxicity tests* should be conducted monthly (i.e., must be within 14 days before or after the date that each toxicity test is initiated) with the laboratory's cultures, using the conditions and procedures outlined in Section 4.8. If this routine is not followed, the performance of individuals from the culture used to start a toxicity test should be evaluated in a reference toxicity test conducted concurrently. This is a requirement for test organisms imported for immediate use in sediment or water-only toxicity tests (see Section 4.8). Additionally, the performance of any cultures that have been recently established using new breeding stock (Section 2.2) should be checked with a reference toxicity test, and the results determined to be acceptable (see Sections 2.3.11 and 4.8) before these cultures are used to provide test organisms.

Environment and Climate Change Canada's Biological Assessment and Standardization Section is investigating the use of positive controls, that are conducted with each test, as an alternative to routine reference toxicant testing. Positive controls are defined as an exposure of test organisms to conditions similar to a negative control (i.e., same number of replicates, number of organisms per replicate, same vessels, etc.) except exposed to a single concentration of a known toxicant. This option would be more feasible and practical for longer term sublethal- and lifecycle-type toxicity tests, such as the 42-day reproduction test with *Hyalella azteca*, and is being considered for non-routine tests conducted infrequently (e.g., one to three time(s) per year).

Typical reference toxicant testing is required to be conducted within two weeks of a test start or can be run concurrently. This is still relevant and important for tests of shorter duration where the design is the same as the environmental/definitive test. However, a positive control run concurrently with every test can have several advantages: economical (reduced effort and resources), reflects a response by organisms sub-sampled from the batch used for testing (especially in the case of imported organisms), and can measure the same endpoint(s) in the same matrix and duration as the definitive test (most notably for sediment and soil toxicity tests).

The choice of toxicant for the positive control can be made using the same criteria as a reference toxicant, but at a single concentration known to elicit a consistent partial response (as compared to reference toxicant tests conducted using multiple concentrations and capture a range of effects; e.g., complete mortality to complete survival). For positive controls, the endpoint would be the mean response (e.g., survival, weight, number of young) subtracted from the mean response in the negative control, divided by the mean negative control response and

---

Section 4.2), but can also be plotted over time (i.e., control performance charts) to assess the ongoing acceptability of the test system and any trends over time that might be indicative of a bias in the test system. The charts are prepared as described for warning charts using reference toxicants (see Section 4.8), and mean values for survival and growth, rather than LC50s or IC50s. As described in Section 4.8, each new mean value should be compared with the established limits of the control performance chart and any trends in the data over time assessed. The USEPA has proposed that in order for laboratories to demonstrate proficiency in carrying out 10-day sediment tests ≤15% of their tests can be below the test validity criteria for survival and growth in control sediment (USEPA, 2011).

multiplied by 100 to provide a percent inhibition or stimulation.

If selecting this option, the positive control response must be defined and include acceptability limits for each endpoint(s). For example, substance X must produce a 40% inhibition (>30 and <50%) and with a coefficient of variation (CV) of ≤40%. Keeping in line with currently required reference toxicant test results, the results of an individual positive control test are not to be used to determine the acceptability of the corresponding test result (i.e., as test validity criteria), but rather can be used to monitor consistency over time (i.e., similar means among positive control tests) and precision over time (i.e., overlapping ranges among positive control tests). Identifying outliers in test organism response or extreme variability in response for individual tests must be used to trigger investigations into potential causes such as culture sensitivity, culture health, environmental/facility conditions, and technician performance. Acceptability limits must be operationally defined at each laboratory and variability limits are fit for purpose. Data obtained from negative controls, positive controls and culture health data should be monitored over time (i.e., by trend analysis) to proactively indicate changes in the organism response.

Cultures should be observed on a frequent and routine basis (e.g., daily or, as a minimum, two or three times per week on non-consecutive days). The estimated number of surviving adults and the production of young in each culture chamber, dates of culture renewals, numbers and age classes of transferred individuals, daily feedings, water quality measurements, etc. should be documented (see Section 9, *Reporting Requirements*).

A summary of the various conditions and procedures that have been used by government laboratories for culturing *H. azteca* is provided in Appendix D. These procedural specifics have presumably worked well in producing test organisms and, unless indicated otherwise in this report, provide useful guidance that may be applied here. A checklist of recommended and required conditions and procedures for culturing *H. azteca* to generate offspring for use in sediment toxicity tests is given in Table 1.

2.3.2 Facilities and Apparatus

H. azteca must be cultured in a controlled temperature laboratory facility. Equipment for temperature control (i.e., incubator, recirculating water bath, or constant temperature room) must be adequate to maintain the temperature within the required limits (Section 2.3.5). The culturing area must be isolated from any testing, sample storage, or sample preparation areas, to avoid contamination from these sources. It must be designed and constructed to prevent contamination of cultures (e.g., elimination of copper or galvanized piping or fixtures that could drip metal-contaminated condensates). The air supply to this area should be designed and operated to prevent entry or recirculation of air from the testing facility or from other portions of the laboratory where contaminants are present.

All equipment, containers, and accessories that might contact the organisms or water within the culturing facility must be clean, rinsed as appropriate, and made of non-toxic materials (e.g., glass, Teflon™, type 316 stainless steel, nylon, Nalgene™, porcelain, polyethylene, polypropylene, fibreglass).

Table 1 Checklist of Recommended and Required Conditions and Procedures for Culturing *Hyalella azteca* for Use in Sediment Toxicity Tests[12]

Source of amphipods	−	existing government, private, or commercial culture; all animals in a test from the same source; species identification confirmed
Acclimation	−	gradually (≤2°C/day) for temperature differences upon arrival
Water source	−	uncontaminated ground, surface, reconstituted, or, if necessary, dechlorinated municipal tap water; reconstituted or natural seawater with salinity ≤15‰ for special needs
Water quality	−	temperature monitored daily; dissolved oxygen monitored at least weekly; pH, hardness, alkalinity, and ammonia measured during 24-h period preceding start of test
Water renewal	−	intermittent-renewal or continuous-flow; ≥1 volume addition/d recommended, 25 to 30%/week (minimum) unless water is recirculated through a filtration system
Temperature	−	23 ± 1°C as daily average, and 23 ± 3°C as instantaneous
Aeration/oxygen	−	aerate gently (recommended); maintain dissolved oxygen at ≥80% saturation
Lighting	−	2000 to 2500 lux adjacent to the water surface; overhead full-spectrum tubes (fluorescent or equivalent, with a broad-spectrum wavelength); photoperiod 16-h light:8-h dark
Substrate	−	medicinal gauze bandage; other choices (e.g., see Appendix D.5) allowed
Feeding	−	various types, quantities, and rates allowed
Age for test	−	14-day survival and growth test: 2- to 9-day old at start of test; should range in age by ≤2 days (must be ≤3 days); 42-day survival, growth and reproduction test: 7- to 9-day old at start of test
Health criteria	−	For in-house cultures: discard batch of organisms intended for use in a test if >20% of young amphipods die or appear stressed during the 48-h period before the test For imported organisms: discard batch of organisms intended for use in a test if >20% of test organisms die or appear stressed during the 24-h period before the test; in addition, if imported for use in the 42-d survival, growth and reproduction test method, organisms (7- to 9-days old) must have a minimum acclimation time of 24 hours prior to use

[12] Conditions and procedures listed apply primarily to known-age cultures (Section 2.3.10), which are commonly maintained in 1- to 2-L beakers or jars, and do not necessarily apply to large or mixed-age stock cultures.

Toxic materials including copper, zinc, brass, galvanized metal, lead, and natural rubber must not come in contact with this apparatus and equipment, or the culture water. Online compressed air to the culturing facility should be filtered as necessary to ensure that it is free of oil and fumes.

2.3.3 Lighting

Overhead full-spectrum lights (fluorescent or equivalent) should illuminate the cultures. The *photoperiod* should be regulated at 16-h light and 8-h dark, and the light intensity adjacent to the water surface in the culture chambers should range within 2000 to 2500 *lux* (MESI, 2010).[13]

2.3.4 Culture Water

Sources of water for culturing *H. azteca* may be an uncontaminated supply of groundwater, surface water, or reconstituted water. Culture water may also be prepared by diluting natural water with a high purity distilled or *deionized water* until a desired hardness is achieved. Acceptable water must allow satisfactory survival, growth, and reproduction of this species. For certain site-specific investigations, the experimental design might require use of water taken from the *site* where sediment is collected. If this or other surface water is used, it should be filtered through a fine-mesh net (e.g., 30 μm) to remove potential predators or competitors. Water that might be contaminated with pathogens may be sterilized by passing it through an ultraviolet sterilizer

Dechlorinated water is not recommended for use as culture or test water, since its quality is often variable and it could contain unacceptably high concentrations of chlorine, chloramines, fluoride, copper, lead, zinc, or other contaminants. Notwithstanding, certain laboratories routinely use dechlorinated municipal water for culturing *H. azteca* and as test water with no apparent problems. If municipal drinking water is used, effective dechlorination[14] must remove any harmful concentration of residual chlorine or chloramines.[15]

If reconstituted fresh water is used for culturing *H. azteca*, the five-salt reconstituted water (SAM-5S) developed at NWRI (Borgmann, 1996, 2002; Borgmann *et al*., 2005b) is recommended (See Section 1.4). The following recipe, which provides reconstituted water with a hardness of 120 to 140 mg $CaCO_3$/L, is taken from Borgmann (1996). It has a higher bromide concentration[16] compared to other recipes

[13] In a recent multi-laboratory method development study, it was determined that laboratories using higher light intensities for culturing (i.e., 2000 to 2500 lux versus the 500 to 1000 lux range recommended in the 1997 version of EPS 1/RM/33) had improved reproduction rates in cultures (MESI, 2010).

[14] Vigorous aeration of the water can be applied to strip out a portion of any residual chlorine gas. This could be followed by use of activated carbon (bone charcoal) filters and perhaps subsequent ultraviolet radiation (Armstrong and Scott, 1974) for removing most of the residual chloramine and other chlorinated organic compounds. Aging the water in an aerated holding tank for one or two days might be of further benefit.

[15] The target value for total residual chlorine, recommended for the protection of freshwater aquatic life, is 0.5 μg/L (CCME, 1999). However, the lowest reliable limit of detection reported is 10 μg/L, suggesting any detection is an indication that aquatic life could potentially be negatively affected.

[16] Successful toxicity testing with *Hyalella azteca* in artificial media has been inconsistent (USEPA, 1994a, 2000; Borgmann, 1996). In 1996, however, Borgmann reported success using the addition of sodium bromide in reconstituted water, and since then, several studies have confirmed that bromide can be used in artificial medium to support more consistent and acceptable survival, growth, and reproduction of *H. azteca* in toxicity tests (Borgmann, 1996, 2002; Ivey *et al*., 2004). In a recent study, Ivey *et al*. (2011) confirmed that survival, biomass and reproduction of *H. azteca* improved in a variety of reconstituted waters with the addition of ≥0.02 mg/L of bromide, a level found in natural waters. Therefore, concentrations of bromide as

commonly used (e.g., USEPA, 1985a, b, 1991c, 1993, 1994a, 2000; EC, 1992b), and in many laboratories this recipe has been found to be suitable for culturing *H. azteca*.[17] To prepare 40 L of SAM-5S reconstituted fresh water, use reagent grade chemicals (anhydrous salts) as follows (Borgmann, 1996):

1. To 100 mL of high purity *distilled*, *deionized* or *reverse osmosis* water in a glass beaker add the following:
   - calcium chloride ($CaCl_2$) – 4.44 g[18]
   - sodium bicarbonate ($NaHCO_3$) – 3.36 g
   - magnesium sulphate ($MgSO_4$) – 1.20 g
   - potassium chloride (KCl) – 149 mg
   - sodium bromide (NaBr) – 41.2 mg
2. Stir the contents of the beaker until all of the salts are dissolved.
3. Place about 20 L of high purity deionized or distilled water in a clean (see Section 3.1) container or carboy.
4. Pour the contents of the beaker (i.e., 100 mL water containing the dissolved salts) into the carboy, ensuring the entire contents of the beaker is transferred (i.e., rinse the beaker with a little distilled or deionized water and add to the carboy), and fill the carboy to 40 L with deionized or distilled water.
5. Aerate the mixture for at least 24 h at room temperature before use.
6. The water quality of the reconstituted water should be approximately the following: hardness, 120 to 140 mg/L as calcium carbonate ($CaCO_3$); alkalinity, 60 to 80 mg/L as $CaCO_3$; *conductivity*, 300 to 500 µS/cm; and pH, 6.5 to 8.5.

The reconstituted water should be aerated for a minimum of 24 h before use to adjust the dissolved oxygen to an acceptable range (see Section 2.3.6) and to stabilize pH. Conductivity, pH, hardness, dissolved oxygen, and alkalinity should be measured in each batch of reconstituted water (USEPA, 1994a, 2000). The reconstituted water may be stored at room temperature (20 ± 3°C) for up to one month if in a clean carboy and capped to prevent contamination (P. Jackman, ALET, Environment Canada, Moncton, NB, personal communication, 2012).

The concentration of salts in the reconstituted water can be adjusted to be similar composition to a receiving water of interest; however, the Ca:Br ratio must be kept constant since these are essential for *H. azteca* and must be present together.

Natural or reconstituted seawater with a salinity of ≤15‰ may be used for culturing *H. azteca* (USEPA, 1994a, 2000). Reconstituted seawater is prepared by adding hypersaline brine, an acceptable formulation of reagent-grade salts, or commercially available dry ocean salts (e.g., Instant Ocean™) to deionized or distilled water or a

low as 20 µg/L in reconstituted water may be sufficient to support *Hyalella* growth and reproduction. In addition, a minimum chloride concentration of 15 mg/L (added as NaCl) is also required to support reproduction (Ivey, *et. al.,* 2016).

[17] Some laboratories have experienced only marginal success in culturing *H. azteca* using reconstituted water, and prefer natural well or surface water for this purpose (G.A. Burton, Jr., Wright State Univ., Dayton, OH, personal communication, 1994).

[18] For $CaCl_2 \cdot 2H_2O$ the amount is 5.83 g.

suitable uncontaminated fresh water, in a quantity sufficient to provide the desired salinity (EC, 1992a, 1997c; USEPA, 1994b).

The characteristics of the water used within a laboratory for culturing *H. azteca* should be reasonably uniform, in order to improve the likelihood of intralaboratory culturing success and to minimize variations in condition and development of cultured organisms. According to USEPA (1994a, 2000), a natural water is considered to be of uniform quality if monthly ranges of the hardness, alkalinity, and specific conductance are less than 10% of their respective averages, and if the monthly range of pH is less than 0.4.

The quality of water in culture chambers should be monitored and recorded routinely. Water temperature should be measured daily, and dissolved oxygen measured at least weekly. Culture water hardness, alkalinity, pH, and ammonia should be measured as frequently as necessary to document water quality. It is recommended that these variables be measured at least quarterly, as well as on the day before the start of a test (USEPA, 1994a, 2000).

Water used for culturing *H. azteca* should be analyzed for nitrite, suspended solids, total dissolved gases, metals, pesticides, and any other contaminants of concern, as frequently as necessary to document water quality (e.g., quarterly). For each analytical method used, the detection limit should be appreciably (e.g., 3 to 10 times) below either (a) the concentration in the water, or (b) the lowest concentration that has been shown to adversely affect the survival, growth, or reproduction of *H. azteca* or other sensitive freshwater animals (EC, 1992b).

The water within culture chambers should be renewed routinely.[19] This may be accomplished manually, or automatically using suitable apparatus and techniques for continuous or *intermittent renewal*. A *water renewal* rate equivalent to ≥1 volume addition/day has been recommended (USEPA, 1994a, 2000), although such a frequent rate of exchange is likely unnecessary. A volume addition of 25 to 30% per week is the minimum exchange allowable (NWRI, 1992) unless water is re-circulated through commercial (aquarium supply) filters.

2.3.5 Temperature

The temperature of the water in culture chambers containing *H. azteca* should be 23 ± 1°C, as a daily average (Table 1). Additionally, the instantaneous temperature of the culture water should be 23 ± 3°C.

2.3.6 Dissolved Oxygen

Water to be used for cultures should be aerated vigorously just before use, to ensure adequate oxygen content and to prevent supersaturation with gases. Dissolved oxygen (DO) should be measured at this time to confirm that a satisfactory value has been obtained (e.g., 90 to 100% saturation).

Cultures should be aerated gently (e.g., 1 bubble/s for each litre of water; Brooke *et al.*, 1993) using filtered, oil-free compressed air. Air to cultures should be dispensed through disposable airline tubing and disposable glass or plastic pipettes or, for large volume cultures, aquarium supply airstones. To ensure that dissolved oxygen is

[19] Continuous recirculation of culture water through commercial aquarium supply filters can also maintain good quality water within cultures (G.A. Burton, Jr., Wright State Univ., Dayton, OH, personal communication, 1994), and such apparatus may be used as an alternative or supplement to water renewal.

adequate to sustain optimum survival and growth of amphipods, it is recommended that DO in cultures be maintained at 80 to 100% saturation.

2.3.7 Culturing Substrate

Various types of substrate have been used successfully for culturing *H. azteca* (see Appendix D.5); the choice is left to the discretion and experience of laboratory personnel. Presoaked medicinal cotton gauze strips (e.g., 5 × 10 cm or 3 cm^2, depending on the size of the culture chamber) are frequently used and are recommended as a suitable substrate (Borgmann *et al.*, 1989; DFO, 1992). The USEPA (1994a, 2000) recommends soaking cotton gauze in water for 24 h before use, and replacing the gauze weekly. Other materials including Nitex® nylon mesh, plastic mesh, silica sand or shredded paper towels (Appendix D.5) may also be used.

2.3.8 Food and Feeding

Various types of food and feeding regimes have been used for culturing *H. azteca* (see Appendix D.6). Success in culturing this species has been achieved using a single ration diet such as commercial fish food flakes (e.g., Nutrafin®, Tetrafin®, TetraMin® or Zeigler® Aquatox Feed) or rabbit chow (Ingersoll and Nelson, 1990; DFO, 1992; NWRI, 1992; Milani *et al.*, 1996), as well as a mixed diet such as filamentous algae, yeast, cereal grass media (e.g., Cerophyll™) and trout chow (USEPA, 1991b; Brooke *et al.*, 1993). In their 1994 method document, the USEPA recommended feeding cultures a yeast-cereal grass media-trout chow (YCT) mixture together with the green algae *Selenastrum capricornutum* and the diatom *Navicula* spp., three times per week (USEPA, 1994a). More recently however, the USEPA (2000) describes several options for feeding cultures, including a mixture of YCT and green algae, or commercial fish food flakes. Studies conducted by Ivey *et. al.* (2016) have also included diatoms (*Thalassiosira weissflogii)* as a dietary supplement. The choice of food type and ration to be used for culturing *H. azteca* is left to the discretion and experience of laboratory personnel. Instructions for the preparation and storage of YCT are given in Section 4.4 and Appendix H. Instructions for the preparation of commercial fish food flakes are given in Section 4.4 and associated footnote (43).

2.3.9 Handling Organisms

Amphipods should be handled as little as possible. When handling is necessary, it should be done gently, carefully, and quickly to minimize stress to the animals. Adult or younger individuals can be transferred between containers using a glass or clear plastic pipette with a polished end and an opening of about 5 to 6 mm in diameter. Transferred organisms should be released below the water surface. When handled, any animals that are dropped, injured, contact dry surfaces, or appear stressed must not be used for testing.

2.3.10 Mixed Age and Known Age Cultures

There are various options for the type, size, and loading densities of culture chambers (see Appendix D.2), as well as the type of culture water used and its method of replacement (Appendix D.3), choice of culturing substrate (Appendix D.5), food type and feeding frequency during culturing (Appendix D.6), and techniques for harvesting young for tests (Appendix D.7). For this test method, such choices are left to the discretion and experience of laboratory personnel; however, the culturing procedures used must produce enough 2- to 9-day old amphipods, ranging in age by ≤3 days (≤2 days is recommended) to start the planned toxicity tests with 14-day sediment and water-only tests, as well as the

reference toxicant tests. For the 42-day survival, growth and reproduction test, culturing procedures must be such that sufficient organisms are available within the age range of 7- to 9-days old. Additionally, the cultured organisms must meet specific performance criteria (see Sections 2.3.11, 4.7, and 4.8).

Laboratories culturing *H. azteca* frequently maintain both *mixed age* and *known age cultures*, and such practice is recommended here. The mixed age culture(s) could represent a mass culture of amphipods of various ages, maintained in one or more aquaria (see Appendix D.2).[20] The known age cultures contain individuals of a particular age class (e.g., <1- to 7-day or 7- to 14-day old) that have been segregated and maintained in a number of aquaria, jars, or other culture chambers (Appendix D.2) until they are used in toxicity tests. Various procedures exist for culturing known age individuals (USEPA, 1994a, 2000), and those that work for laboratory personnel may be used to provide test organisms. In each of these procedures, the water in the culture chambers is changed routinely, and thus known age animals are obtained.

Following is a procedure (Hamr *et al.*, 1994; Milani *et al.*, 1996) for generating <1- to 7-day old *H. azteca*, and for holding them for a subsequent 2-day period (i.e., until they are 2- to 9-days old) before their use in a 14-day sediment toxicity test. Hamr *et al.* (1994) provide a rationale and experimental data that support this choice of age of test organisms and a 14-day test duration. The procedure for obtaining animals within this age range has been modified from Borgmann *et al.* (1989). This procedure is recommended as one of several ways of obtaining adequate numbers of 2- to 9-day old organisms (at the time they are used in the sediment toxicity tests).

Approximately 150 adult amphipods that are ≥30-d old and ideally in amplexus are placed in 20-L culture chambers, each containing about 15 L of culture water. Each chamber contains pieces of cotton gauze. The cotton gauze should be presoaked in water for 24 h before its initial use, and can be used for up to three weeks before replacement. A combined yeast-cereal grass media-trout chow (YCT) preparation (see Appendix H) is provided daily to each culture chamber. Once a week, the test organisms are isolated from the gauze and the culture water by pouring the contents of the culture chamber(s) through 500 µm and 250 µm sieves. The animals retained by the 500 µm sieve are returned to the culture chambers, which are replenished with at least 30% fresh water and food. The animals retained by the 250 µm sieve are <1- to 7-days old. These animals are rinsed into a translucent white plastic pan placed on a light table and are counted. The young amphipods from each 20-L culture chamber are then transferred, using a pipette, into a 1-L beaker, which contains 750 mL of fresh culture water. Density of amphipods in beakers should not exceed 1 animal per 10-mL volume of culture water (i.e., 75 animals/beaker) to avoid growth inhibition (K. Day, NWRI, Burlington, ON; unpublished data). Presoaked cotton gauze should be placed in the beaker to provide a substrate for the animals. Organisms in each 1-L beaker are fed 10 mL of YCT daily, and for the 14-day exposures, are held for 2 days before starting a toxicity test. Accordingly, the animals are 2- to 9-days old at the start of the test. Each culture chamber containing 150 animals will produce 100 to 150 young per week, on average.

[20] Some laboratories have reported improved growth and reproduction in mass cultures where aquaria are not maintained in pristine conditions and algae is encouraged to grow.

For the 14-day exposures, a narrower age-range of organisms used to start a test is required (i.e., organisms must range in age by ≤3 days; however, a narrower age-range of ≤2 days is recommended) to reduce the potential for variability in the results, especially since final dry weight (as an indication of growth) is measured as a primary test endpoint. For the 42-day exposure the narrower and older age range (i.e., 7 to 9 days old) is required in order to ensure adequate growth in the control sediment to reproduce at approximately days 24 to 28. Isolation of about 1500 (750 pairs) adults in amplexus will provide about 800 newborn amphipods in 24 h, and requires about six person-hours of time (USEPA 1994a, 2000).

Records should be kept on the number of surviving adults, number of breeding pairs, and the number of young produced and their survival. Records should also be kept on the age of brood organisms, and on the frequency of restarting cultures. This information can be used to develop performance charts, which are useful in determining whether cultures are maintaining a vigorous reproductive rate indicative of culture health. Some of the adult amphipods can be expected to die in the culture chambers, but excess mortality should be cause for concern. A decrease in reproductive rate could be caused by a change in water or food quality, or by deteriorating health of the brood stock. Culture performance is affected by the age of adults and can be cyclical. Adult females will continue to reproduce for several months; however, fertility gradually decreases after about three months (USEPA, 1994a, 2000).

2.3.11 Health Criteria

Amphipods in the cultures should be checked three times per week (e.g., Monday, Wednesday, Friday) as a minimum, and preferably daily. Individuals that appear unhealthy (e.g., discoloured or otherwise stressed), inactive, or dead when gently prodded must not be used for testing. If more than 20% of the amphipods in a *known age* culture chamber appear dead or inactive during the 48-h period preceding the start of the test, the entire group in the container must be discarded (USEPA, 1994a, 2000).

Ideally, a reference toxicity test should be performed in conjunction with each sediment and water-only toxicity test. Laboratories routinely undertaking sediment toxicity tests using *H. azteca* may choose to conduct reference toxicity tests once each month according to a regular schedule; however, reference toxicity tests must be conducted within 14 days before or after the start of each toxicity test. All tests with reference toxicants should be performed using the conditions and procedures outlined in Section 4.8. ECCC is investigating the use of positive controls for the 42 day reproduction test (Section 2.3.1) to allow for matrix and endpoint matched responses to reference toxicants. Laboratories may pursue this option when conducting the 42 day test only. Test-related criteria used to judge the health and sensitivity of the culture are given in Sections 4.7 and 4.8.

Biochemical measurements such as the lipid content of cultured amphipods, or the average number of young produced in a week by each adult in a culture, might provide useful information on the health of the cultures (USEPA, 1994a, 2000). Ongoing records of these or other indices of the condition of cultures will likely prove useful and are encouraged. No specific health criteria have been developed as yet with respect to physiological measurements, although they could be applied in the future.

Section 3

Test System

3.1 Facilities and Apparatus

Tests may be performed in a water bath, environmental chamber, or equivalent facility having acceptable temperature and lighting (see Section 3.2) control. The test facility must maintain the daily mean temperature of all sediment and water in test vessels at 23 ± 1°C for the 14-day survival and growth tests (see Section 4.2) and at 23 ± 2°C for the 42-day survival, growth and reproduction test (see Section 8.1). The facility should be well ventilated to prevent exposure of personnel to harmful fumes, and isolated from physical disturbances or any contaminants that might affect the test organisms. The area used to manipulate sediment in preparation for tests should also be properly ventilated.

The test facility must be isolated from the area where amphipods are cultured, to avoid potential contamination of cultures. Additionally, the test facility should be removed from places where samples are stored or prepared, to prevent the possibility of contamination of test vessels and their contents from these sources. The ventilation system should be designed, inspected, and operated to prevent air within the testing facility from contaminating culture facilities. Return air from sample handling and storage facilities or those where chemicals are processed or tested should not be circulated to the area of the laboratory where tests are conducted. Any construction materials that might contact the organisms, water, or test vessels within this facility must be non-toxic (see Section 2.3.2).

Compressed air used within the test facility for aerating water must be free of oil and fumes. Oil-free air pumps should be used wherever possible. Any oil or particulate in the air supply should be removed by online filters, which are replaced as required to ensure their effectiveness.

Equipment and supplies that contact water, sediment, or test vessels should be chosen to minimize sorption of chemicals. Borosilicate glass, nylon, high-density polyethylene, polycarbonate, fluorocarbon plastics, and type 316 stainless steel should be used whenever possible to minimize chemical sorption and leaching. The use of apparatus and supplies made of toxic substances (see Section 2.3.2) must be avoided.

The 14-day survival and growth sediment toxicity tests must be performed as a static test (i.e., no renewal of overlying water). However, if the test water overlying any *reference sediment* is fouled by high levels of ammonia, and/or deteriorates due to low DO levels or pH levels outside the tolerance range of *Hyalella azteca*, and the objective of the test is to exclude the effects of these confounding factors from the evaluation of the toxicant, the test must be conducted as, or shift to, a static-renewal exposure, which requires a minimum of 3 times-weekly renewal (i.e., ≥6 times during the test) of the overlying water in test vessels (see Section 4). Ammonia is present in freshwater sediment as a result of the natural processes of decomposition of the organic matter incorporated into the sediment, as well as anthropogenic sources. Rising levels of ammonia in overlying water can also be caused by excess or uneaten food. In addition, oxidation of sulphides in sediment samples can lead to the reduction of pH during testing to extremely low levels (e.g. pH 4), resulting

in complete mortality of *Hyalella* in uncontaminated reference sediments (Borgmann and Norwood, 1999). *H. azteca* can survive exposure to extremely low levels of DO; however, these levels can affect their growth (see Section 1.4). High levels of ammonia as well as low levels of pH and DO (i.e., those that impact the survival and/or growth of the *Hyalella*) occurring in reference sediments (presumably clean) collected during each field collection might confound the interpretation of the sediment toxicity test results. This revised and updated version of Report EPS 1/RM/33, therefore proposes ammonia concentrations (>0.2 mg/L unionized NH_3-N),[21] pH levels (<6.0 or >8.0), and DO levels (<40%)[22] in the *test water* overlying reference sediments that should be used as triggers for conducting the test as a static-renewal exposure rather than a static one, if the objective of the test is to assess the toxic effects due to substances or materials without the effects of these confounding factors.

Water-only toxicity tests must be performed as a static-renewal exposure, which requires a minimum of 3 times-weekly renewal (i.e., ≥6 times during the test) of test solutions (see Section 7.5.4).

The 42-day survival, growth and reproduction test must be performed as a static-renewal exposure with renewal of overlying water on Days 14, 28 and 35 (see Section 8.11).

Overlying *water renewals* may be done manually (see Sections 4.1 and 7.5.4) or automatically (USEPA, 2000). Apparatus used for the automated delivery and renewal of overlying water in test vessels may be one of several designs (e.g., Maki, 1977; Benoit *et al.*, 1993; Zumwalt *et al.*, 1994). A suitable apparatus would be one that enables the timed and *intermittent renewal* of the overlying water in each test vessel at a rate of two volume additions/day (USEPA, 1994a, 2000).

[21] A >0.2 mg/L unionized NH_3-N is recommended herein as the level of ammonia that might trigger the initiation of a static-renewal exposure (as opposed to the standard static exposure), depending on the study objectives. This level is based on the results of a study carried out by Environment Canada's Atlantic Laboratory for Environmental Testing (ALET, Moncton, NB) to determine the effect of confounding factors (i.e., ammonia) on sediment toxicity tests. In this study, silica sand spiked with ammonium chloride was tested with overlying water also spiked with ammonium chloride and, in a second experiment, with clean overlying water (i.e., not spiked). The 14-day IC25s for *H. azteca* growth based on ammonia levels in the pore water ranged from 0.2 to 1 mg/L unionized NH_3-N, and based on ammonia levels in the overlying water, the IC25s ranged from 0.2 to 0.4 mg/L unionized NH_3-N (Jackman and Doe, 2000). Data from various literature sources was also compiled to assist in the determination of a level of ammonia that could be deemed as detrimental to *Hyalella* and might therefore trigger a test to be carried out as static-renewal; however, many of the studies were not relevant (water-only exposures, different test durations, etc.) and none included *sublethal* endpoints. The two most relevant studies (Besser *et al*., 1998 and Whiteman *et al*., 1996) were 4-day sediment tests using *H. azteca*, and the LC50s reported were in keeping with the data produced by ALET in their ammonia study. Results produced by both Whiteman *et al*. and ALET did show that the LC50 for ammonia in pore water was approximately 10-times higher than that for overlying water. Whitman *et al*. (1996) attributed this difference to avoidance of the sediment by *H. azteca*. This data also indicates that the concentration of ammonia in the overlying water may be more relevant to *H. azteca* toxicity than that in the pore water, and that a higher concentration of ammonia in the pore water (i.e., >0.2 mg/L unionized NH_3-N) may be necessary to actually induce ammonia toxicity. It has been well documented that many factors affect the toxicity of ammonia to *H. azteca*, further confounding the issue. These include: temperature, pH, hardness, and the ionic composition (i.e., sodium, potassium, bromide, and chloride) of the medium in which the *H. azteca* are being tested (Borgmann, 1994, 1996; Ankley, 1995; Borgmann and Borgmann, 1997; Wang *et al*., 2008).

[22] The pH and DO levels recommended herein are based on literature data on the pH and DO tolerance of *H. azteca* (see Section 1.4).

The USEPA (2000) provides useful guidance on the design of this and other suitable systems for the automated renewal (at a rate of two volume additions/day) of overlying water.

A supply of disposable glass pipettes and aquarium supply airline tubing is required for delivery of a continuous (gentle) flow of air to each test vessel. Stainless-steel (rather than brass) gang valves are recommended for regulating air flow. In addition, a supply of suitably sized watchglasses or lids is needed for covering individual test vessels.

The test facility must have the basic instruments required to monitor the quality (e.g., temperature, pH, dissolved oxygen, conductivity) of the test water and pore water. Additionally, the laboratory should be equipped to facilitate prompt and accurate analysis, with acceptable limits of detection, of such variables as hardness, alkalinity, ammonia, and (in instances where dechlorinated municipal water is used as culture or test water) residual chlorine.

All test vessels, equipment, and supplies that might contact sediment or test water must be clean and rinsed with test water, deionized water, or distilled water, before its use. All nondisposable materials should be washed after use. The following cleaning procedure (USEPA, 1994a, 2000) is recommended.

1. Soak in tap water for 15 minutes, then scrub with detergent or clean in an automatic dishwasher.

2. Rinse twice with tap water.

3. Rinse carefully with fresh, dilute (10%, v:v[23]) nitric (HNO_3) or hydrochloric acid (HCl) to remove scale, metals, and bases.

4. Rinse twice with deionized water.

5. Rinse once with full-strength, pesticide-grade acetone to remove organic compounds (use a fume hood or canopy).

6. Rinse three times with high-quality deionized water.

Before toxicity tests are undertaken in a new test facility, a minimum of five 96-h water-only reference toxicity tests and a minimum of five 14-day *survival-and-growth* tests using control sediment (for sediment tests), or control/dilution water (for water-only tests), and different batches of test organisms should be undertaken to confirm that acceptable performance of *Hyalella azteca* can be achieved using the new facility and the culturing and test conditions and procedures specified in this report (see Sections 2.3.1, 4, 7 and 8). Each test with reference toxicant, control sediment, or control/dilution water should be performed using a different *batch* of cultured organisms. Data from these preliminary tests should be compared by calculating and appraising the magnitude of the coefficient of variation for the respective series of tests and endpoint values.

3.2 Lighting

All test vessels should receive full-spectrum (e.g., fluorescent or equivalent) illumination from directly overhead, at an intensity sufficient to provide 500 to 1000 *lux* adjacent to the surface of water or overlying water in

[23] To prepare a 10% solution of acid, carefully add 10 mL of concentrated acid to 90 mL of deionized water.

test vessels. Illumination should be as uniform as possible for all test vessels. Photoperiod should be regulated at 16-h light and 8-h dark.

3.3 Test Vessels

Glass beakers or glass jars must be used as test vessels. When conducting the 14-day survival and growth tests, high-form glass vessels with a capacity of 300 mL and an inner diameter of ≥7 cm are recommended (USEPA, 1994a, 2000; ASTM, 1995a, 2010; EC, 1997a). When conducting the 42-day survival, growth and reproduction test, wide-mouth 1 L glass beakers or glass jars (e.g., 1 L mason jars approximately 4 inch diameter) are required.

Each beaker or jar must be cleaned thoroughly before and after use (Section 3.1), and rinsed well with test water immediately before use. Covers should be used for all tests, especially if *test sediments* contain detectable volatile gases. Suitable covers include clean watch glasses, or glass or plastic lids. Covers must be transparent, and allow for some air exchange (e.g., petri-dish with holes for aeration and feeding).

3.4 Test and Control/Dilution Water

Depending on the test design and intent (see Sections 5 and 6), test water (i.e., water overlying sediment in the test) and control/dilution water (i.e., water used to prepare dilutions of test chemicals and as control water in *water-only* exposures) may be either an uncontaminated supply of natural fresh or estuarine water, or reconstituted water. The water supply used as test or control/dilution water is frequently the same as that used for culturing *H. azteca* (see Section 2.3.4), although it may come from another source. For instance, the use of site water, or clean water adjusted to the hardness of water at a collection site, might prove a good choice (see Section 5.4). However, caution is recommended when considering the use of reconstituted or other overlying waters (e.g., site water) for the 42-day survival, growth and reproduction test, as method development studies showed some variability in achieving the test validity criteria when reconstituted water was used as overlying water. Advanced testing would be required to demonstrate that the 42-d test can be successfully conducted (i.e., meet test validity criteria) using reconstituted or other site waters during the test.

The quality of test water and that used as control/dilution water is extremely important; this water must have been demonstrated to allow acceptable survival, growth and reproduction of test organisms in tests with control sediment (see Section 4.7) before it is used in toxicity tests.[24] However, control/dilution water (including site, reconstituted or other waters) used as overlying water successfully for the 14-day survival and growth test must not automatically be assumed to be acceptable for the longer term 42-day exposure. Pre-testing should be conducted to ensure the test water selected can meet the validity criteria for the longer-term reproduction test.

In addition, when used in the 42-day exposure, the test water must contain a minimum bromide (Br) concentration of

[24] The USEPA (2011) has recently proposed guidance for laboratories to demonstrate the suitability of the food and water being used for *Hyalella* sediment tests, in a performance-based demonstration. A 14-day test may be conducted with test vessels containing a thin layer of quartz sand and the test and/or control/dilution water and food normally used for definitive testing. If the test validity criteria are met (i.e., ≥80% survival and final weight of ≥0.10 mg/organism), then the food and water being used may be considered adequate.

0.02 mg/L (added as NaBr) and minimum chloride (Cl) concentration of 15 mg/L (added as NaCl) (Ivey *et al.,* 2016). Most commercial analytical labs do not routinely achieve a Method Detection Limit (MDL) of 0.02 mg/L for Br. Therefore, in the event Br concentration cannot be confirmed in the overlying water to be used in testing, it must be augmented prior to use to ensure a minimum Br concentration of 0.02 mg/L (also see Section 8.4).

When site water is used as test water, a second set of *controls* must be prepared using a supply (source) of laboratory water shown previously by the testing laboratory to routinely enable valid test results for the applicable test duration (i.e., in a 14-day test for survival and growth or in a 42-day test for survival, growth and reproduction of *Hyalella* [25]). Unless the testing laboratory is importing test organisms rather than maintaining cultures of *Hyalella* at their facility, the laboratory water in which amphipods were cultured must be used for this purpose. In instances where the testing laboratory imports the test organisms, an alternate source of uncontaminated laboratory water shown previously by that laboratory to enable valid test results may be used as the second *control* solution (e.g., reconstituted [SAM-5S] water). Guidance for preparing reconstituted fresh water or estuarine water (salinity ≤15‰) is provided in Section 2.3.4.

Test and control/dilution water must be adjusted to the test temperature (23 ± 1°C when conducting the 14-day tests and 23 ± 2 °C when conducting the 42-day test) before use. The dissolved oxygen content of the water should be 90 to 100% of the air saturation value at this temperature. As necessary, the required volume of water should be aerated vigorously (oil-free compressed air passed through airstones) immediately before use, and its dissolved oxygen content checked to confirm that 90 to 100% saturation has been achieved. The pH of the water should be measured and stable before use.

3.5 Control Sediment

Each sediment toxicity test must include an experimental control, with a minimum of five *replicate* vessels (i.e., beakers or jars), containing control sediment for the 14-day survival and growth tests, and a minimum of eight replicate vessels containing control sediment for the 42-day survival, growth and reproduction test. A control sediment is a sediment that is essentially free of any contaminants that could adversely affect the survival, growth, or behaviour of *H. azteca* during the test. The use of control sediment provides a measure of test acceptability (i.e., mean survival at test end must be ≥80%), evidence of the health and behaviour of the test organisms, and a basis for interpreting data derived from the test sediments.

Control sediment may be either natural sediment taken from a collection site removed from known sources of contaminants and shown previously to enable acceptable survival, growth and reproduction of

[25] If the intent of the test is to measure the extent to which a particular receiving water might modify the toxicity of the test sediment or test material due to its physicochemical characteristics (e.g., hardness, pH, *turbidity*, humic or fulvic acid content) and/or the presence of other contaminants, the investigator might choose to use the *upstream water* as test water (i.e., overlying water or control/dilution water). A comparison of controls for this water with those for the controls held in laboratory water will identify toxic effects that might be contributed by the upstream water. A clearer understanding of the differing influence of each type of test or control/dilution water on the toxicity of the test material can be achieved by undertaking a side-by-side comparison of toxic effects using each type of water to prepare test treatments.

H. azteca under the conditions of the test,[26] or formulated sediment.[27] The selection of an appropriate control sediment depends on considerations such as the study design, physicochemical characteristics of the test sediment(s), and the availability of suitable *clean sediment* having the desired properties. There should also be evidence that the sediment provides consistent and acceptable biological endpoints using *H. azteca* and this test method. While many clean, natural sediments have been used as control sediment in toxicity tests with *H. azteca*, the use of formulated control sediment offers a more consistent, standardized approach and one that numerous researchers are now actively pursuing (Smith *et al.*, 1992b; Dwyer *et al.*, 1993; Suedel and Rodgers, 1994a, b; USEPA, 1994a, 2000; Suedel *et al.*, 1996).

As indicated for the selection of overlying water, control sediment (including formulated sediments) used successfully for the 14-day survival and growth test must not automatically be assumed to be acceptable for the longer term 42-day exposure. Pre-testing must be conducted to ensure the sediment (natural or formulated) can meet the validity criteria for the longer-term reproduction test.

There are a number of acceptable approaches to preparing and conditioning artificial sediment (USEPA, 2000; ISO, 2010). In general, the following attributes should be considered when selecting a formulation for a control or test sediment:

- should support the survival, growth or reproduction of a variety of benthic organisms;
- should provide consistent acceptable biological endpoints for a variety of species;
- should be comprised of standard constituents that are readily available to test laboratories; and
- should be free from concentrations of contaminants that might cause adverse effects to test organisms.

The following artificial sediment can be used as a control for fresh water sediment tests or as a clean material to be spiked with a test chemical (ISO, 2010). Mix the following dry ingredients in the proportions (% mass fraction) given:

40% Silica sand (0.1 mm to 0.4 mm)[28]

[26] Control sediment should be one that readily supports test organisms, but is not so rich that the organisms' growth and reproduction is better than might be typical of most field sediments. If a laboratory's control sediment is too "nutritious," there is a risk that uncontaminated, less "nutritious" test sediments would appear toxic. To evaluate the quality of a control sediment, laboratories should conduct a study that includes both their regular control sediment and a quartz sand control. If growth in the regular control sediment is greater than that in a quartz sand control by >20%, then the control sediment might be considered unacceptable. The USEPA (2011) is recommending that laboratories find and use control sediments that meet these criteria (i.e., <20% difference in *H. azteca* growth when compared to a quartz sand control).

[27] Formulated sediment is also described as reconstituted, artificial, or synthetic sediment. It is typically prepared using sand, silt, clay, and non-toxic organic constituents obtained from commercial sources, and is hydrated with reconstituted or natural water. Formulated sediment can be prepared to match different natural sediments with respect to particle size distribution, organic carbon content, pH, cation exchange capacity, etc. (Suedel and Rodgers, 1994b; Milani *et al.*, 1996). Alternatively, one or more recipes can be used to prepare standardized control sediment(s) for routine use in freshwater sediment toxicity tests with *H. azteca* or other infaunal species (Suedel *et al.*, 1996).

[28] Silica sand contains crystalline silica, which has been prescribed as a *designated substance* under the Ontario Occupational Health and Safety Act. A designated substance is defined as one to which exposure of a worker is prohibited, regulated, restricted, limited, or controlled. The Material Safety Data Sheet for silica sand should be obtained and reviewed prior to its use, and personnel should take the appropriate precautions

30%	Silica sand (W4, mean particle size 0.063 mm)[28]
20%	Al_2O_3
4.5%	Fe_2O_3
4%	Peat (decomposed peat from a raised bog, untreated; finely ground and <1 mm sieved)
1%	$CaCO_3$
0.5%	Dolomite (Clay)

This recipe is based on the artificial sediment recommended by ISO for the "Determination of the toxic effect of sediment and soil samples on growth, fertility and reproduction of *Caenorhabditis elegans* (Nematoda)" (ISO, 2010) and in the ISO 16303 method for sediment toxicity testing using *H. azteca* (ISO, 2013).

for protection to prevent inhalation of and contact with this ingredient.

Section 4

Universal Test Procedures – 14 Day Survival and Growth

General procedures and conditions herein apply to each of the described toxicity tests for samples of sediment, particulate waste, or chemical, and to reference toxicity tests. More specific procedures for conducting 14-day survival and growth tests with field-collected samples of sediment or other similar particulate material (e.g., sludge, dewatered mine tailings, drilling mud residue) are provided in Section 5. Guidance and specific procedures for conducting 14-day survival and growth tests with sediment spiked experimentally with chemical(s), contaminated sediment, or particulate waste are given in Section 6. Procedures for carrying out the 14-day water-only toxicity test for testing samples of wastewater, receiving water, or chemicals are described in Section 7. Specific procedures for carrying out the longer-term 42-day survival, growth and reproduction test with field-collected samples of sediment, sediment spiked experimentally with chemical(s), or particulate wastes are given in Section 8.

All aspects of the test system described in Section 3 must be incorporated into these universal test procedures. The summary checklist of recommended and required test conditions and procedures in Table 2 describes not only universal procedures for the 14-day survival and growth tests, but also those for testing specific types of test substances or materials.

Universal procedures are described in this section for performing a 14-day sediment toxicity test in a static system or static-renewal system, depending on the quality and stability of the test water overlying the reference sediment(s), and the objectives of the test (see Sections 1.1 and 3.1). The test must normally be conducted as a static exposure in which the overlying water is not renewed during the test (except for replacing losses due to evaporation), and is aerated continuously. If, however, the test water overlying sediment collected from any reference site becomes fouled or deteriorates due to rising levels of ammonia (i.e., > 0.2 mg/L unionized NH_3-N), pH levels drifting outside the tolerance range of *Hyalella azteca* (i.e., <6.0 or >8.0), and/or extremely low DO levels (i.e., <40%) at any time during the test, and the objective of the test requires that these confounding factors be excluded as part of the measurement of the total effects of the sample, the test must be conducted (i.e., started) or continued[29] as a static-renewal exposure. In the static-renewal exposure, the overlying water in all vessels (i.e., reference sediment, test sediment, and laboratory control sediment) must be renewed a minimum of 3 times weekly on non-consecutive days (i.e., 6 times during the test) at a rate of 2 volume additions in 24 hours, and test vessels are aerated.[30] If resources and study objectives permit, static and static-

[29] A toxicity test initiated as a static exposure can be shifted to a static-renewal exposure based on water quality measurements (ammonia, pH, DO levels in water overlying any reference sediment) taken throughout the test and depending on the test objectives (see Section 3.1).

[30] Similar results are apparently obtained by static and static-renewal tests, when performed according to the procedures defined herein. The performance of these two test options was compared in side-by-side tests using samples of field-collected or contaminant-spiked sediment (Milani *et al.*, 1996). Results for interlaboratory tests with *H. azteca* indicated that test

renewal exposures can be run in parallel in order to determine the effects of the confounding factors on the toxicity of the sample.Special situations might require more frequent renewal of overlying water (e.g., daily-renewal, where the overlying water is renewed at a rate of two volume additions every day). A daily-renewal exposure may be used if (and only if) the quality of the water overlying a reference sediment is suspected of being highly unstable (i.e., the ammonia, pH and/or DO levels in the overlying water continue to shift beyond the acceptable ranges, described above, on a daily basis) and the objectives of the test are to eliminate the effects of deteriorating overlying water quality.[31]

For all 14-day survival and growth tests, the amphipods are fed either an aqueous suspension of YCT, ground commercial fish food flakes, or a 1:1 combination of YCT and ground fish flakes, either daily, or three times per week on non-consecutive days (see Section 4.4). Biological endpoints measured in the 14-day test method are survival and dry weight at test end.

4.1 Beginning the Test

Each test vessel (see Section 3.3) placed within the test facility must be clearly coded or labelled to enable identification of the sample or its concentration. The date and time when the test is started must be recorded, either directly on the labels or on separate data sheets dedicated to the test. The test vessels should be positioned for ease while making observations and measurements. Treatments should be positioned randomly within the test facility (USEPA, 1994a, 2000).

The day that amphipods are initially exposed to samples of test materials or substances is designated Day 0. On the day preceding the start of the test (i.e., Day –1),[32] each sample or subsample of test sediment or similar particulate material, including control and reference sediment, should be mixed thoroughly[33] (see Sections 5.3 and 6.2) to provide a homogeneous mixture consistent in colour, texture, and water content. Quantitative measures of homogeneity might include particle size analysis, total organic carbon, percent moisture, and concentration of specific chemicals.

precision and sensitivity were similar using either system (Milani *et al.*, 1996). Interlaboratory coefficients of variation (CVs) for grand means (all laboratories) of the 14-day survival data for each of 4 samples of field-collected sediment ranged from 3.6 to 19.6% using the static system, and from 2.5 to 11.0% using the static-renewal system. Data for growth were more variable in both systems, with CVs for dry weight of amphipods at test end ranging from 28.4 to 48.8% using the static system and from 26.0 to 35.7% using the static-renewal system. The inclusion of an option for a static-renewal exposure herein, depending on the objectives of the test, is to allow for the testing of sediments that result in overlying water quality conditions (ammonia, pH, or DO) that are not favourable for the survival of *Hyalella.*

[31] If a daily-renewal exposure is being used, water overlying the sediment in each test vessel should be renewed on the day preceding the test (Day –1) as well as throughout the test at a rate of two volume additions per day (USEPA, 1994a, 2000). In daily-renewal tests, aeration of the overlying water is not normally required. The dissolved oxygen concentration in the overlying water will not be below 40% saturation due to its daily renewal, unless the sediment sample used in the test has an unusually high oxygen demand. If at any time during a daily-renewal test the dissolved oxygen is below 40% saturation in one or more test vessels, the overlying water in all test vessels including the controls should be aerated as described in Section 4.3 (USEPA, 1994a, 2000).

[32] In some cases, longer equilibration times (e.g., up to seven days prior to testing) might be necessary depending on the characteristics of the site sediment and/or the study objectives.

[33] Any liquid that has separated from the sample during transport and/or storage must be remixed within the sample.

Table 2 Checklist of Recommended and Required Conditions and Procedures for Conducting 14-day Survival and Growth Sediment Toxicity Tests Using *Hyalella azteca*

Universal		
Test type	−	14-day whole sediment toxicity test; normally no renewal (static test), optional static-renewal of overlying water triggered by ammonia (>0.2 mg/L unionized NH_3-N), pH (<6.0 or >8.0), and/or DO (<40%) of test water overlying reference sediment(s) and test objectives
Water renewal	−	normally no renewal of overlying water during test except for replacement of losses due to evaporation; if static-renewal, overlying water is replaced ≥3X weekly on non-consecutive days (i.e., ≥6X during the test) at a rate of 2 volume additions in 24 hours
Test (overlying) water	−	culture water or other clean ground or surface water; site water; water adjusted to hardness of site water; reconstituted fresh water for a higher degree of standardization; natural or reconstituted seawater with salinity ≤15‰ for tests with estuarine sediment; dissolved oxygen, 90 to 100% saturation when used as overlying water in test
Acclimation	−	if test water is different from culture water, acclimation of organisms to test water is recommended; acclimation should be conducted on the day preceding the start of the test (Day -1); for estuarine sediment, organisms should be gradually acclimated to test water with a salinity similar to the test sediment pore water
Control sediment	−	sample of clean sediment that is used to assess the performance of the test organisms and the acceptability of the test; either natural or formulated sediment
Amphipods	−	removed from known age culture as <1- to 7-d old individuals and fed daily; test organisms 2- to 9-d old, and ranging by ≤3d (recommended ≤2d) at start of test; 10 animals/test vessel
Test vessel	−	glass beaker or glass jar; recommend ≥7 cm inner diameter; recommend 300-mL high-form glass beaker or jar; normally covered
Volume of wet sediment	−	recommend 100 mL; optional, ≥55 mL; must be ≥2 cm depth
Volume of test water	−	recommend 175 mL; optional, volume resulting in a sediment:water ratio of 1:4 (e.g., 55 mL sediment and 220 mL test water, or 100 mL sediment and 400 mL test water)

Number of replicates	−	must be ≥5 *replicate samples*, each a discrete (i.e., different) sample from the same *sampling station*; must be ≥5 replicates (i.e., replicate vessels) for multi-concentration tests (e.g., spiked sediment) and control sediment
Temperature	−	daily average, 23 ± 1°C; instantaneous, 23 ± 3°C
Lighting	−	overhead full-spectrum (fluorescent or equivalent); 500 to 1000 lux; 16-h light:8-h dark
Aeration	−	continuous and minimal (e.g., 2 to 3 bubbles/s, each test vessel)
Feeding	−	aqueous suspension of YCT, ground commercial fish food flakes (e.g., Nutrafin®, Tetrafin®, TetraMin® or Zeigler® Aquatox Feed), or a 1:1 combination of YCT and ground fish flakes, fed daily or three times per week (non-consecutive days); 2.7 mg solids, dry weight (or equivalent) added daily to each test vessel if daily feeding; 6.3 mg dry solids (or equivalent) added each feeding to each test vessel if fed three times per week only
Observations	−	optional: numbers of amphipods in each vessel seen emerged from sediment, and their behaviour (daily or less frequently)
Measurements of overlying water	−	≥3 times/week: DO and temperature for each *treatment* as well as ammonia and pH for each reference sediment; start and end of test: pH, conductivity and ammonia for each treatment; recommend hardness and/or alkalinity at start and end; salinity where appropriate
Endpoints	−	significantly lower survival and final dry weight than in control or reference treatments (based on mean percent survival and mean dry weight, each treatment); 14-d LC50 for multi-concentration test, where appropriate; *ICp* for weight, where appropriate
Test validity	−	invalid if mean 14-day survival in control sediment <80%; invalid if average dry weight for replicate control groups at test end is <0.1 mg/amphipod

Field-collected Sediment or Similar Particulate Material

Transport and storage	−	if sample >7°C, cool to 7°C (ice or frozen gel packs); transport in dark at 1 to 7°C (preferably 4 ± 2°C); store in dark at 4 ± 2°C; test should start within two weeks and must start within six weeks
Reference sediment	−	collected from one or more sampling stations for tests with field-collected sediment; taken from sites presumed to be clean but in the general vicinity of sites where test sediments are collected (i.e., same body of water); frequently selected for use in the toxicity test because of its physicochemical similarity (e.g., particle size and/or organic carbon content) to the test sediments

Sample characterization	−	at least particle size analysis (percent sand, silt, and clay), total organic carbon; other parameters (depending on study objectives) could include: percent water, pore water pH, and pore water ammonia; pore water and sediment sulphide, redox, total inorganic carbon, total volatile solids, biochemical oxygen demand, chemical oxygen demand, cation exchange capacity, total and dissolved metals in sediment and porewater, synthetic organic compounds, oil and grease, petroleum hydrocarbons, *Simultaneously Extracted Metals (SEM)/Acid-Volatile Sulfide (AVS)* and pore water analyses for various physicochemical characteristics)
Preparation of sample	−	only if necessary, remove debris and indigenous macro-organisms using forceps; homogenize sample (including any separated liquid) before the test; if necessary, remove smaller macro-organisms by pressing through fine-mesh sieve (e.g., 0.25 to 0.5 mm), or pass through fine-mesh sieve using liquid that separated from sample during transit and storage and remix this liquid with the sieved sample
Spiked Sediment		
Characterization of chemical(s)	−	information required on stability, water solubility, vapour pressure, purity, and biodegradability should be known for added chemicals spiked into control sediment
Solvent	−	test water is the preferred solvent; if an organic solvent is used, the test must include a solvent control
Preparation of mixtures	−	procedure depends on test design and objectives; might include one or more chemical concentrations mixed in control or test sediment, or specific chemical concentrations added to the test water overlying control sediment; chemical/sediment mixtures may be prepared manually or by mechanical agitation as slurries
Concentration of chemical(s) added	−	normally measure at beginning and end of test, in high, medium, and low strengths as a minimum
Test and dilution water	−	use reconstituted water if a high degree of standardization is required

Immediately following mixing, replicate volumes of the sample should be transferred to the test vessels. Two options for sediment to water ratios are recommended. For the standard 1:1.75, 100-mL volumes of the sediment should be used. The second option should only be used when larger volumes of water are needed for chemical analysis of the overlying water. This is a 1:4 sediment to water ratio, in which a minimum of 55 mL of sediment at a minimum of 2 cm in depth is required (e.g., 55 mL of sediment with 220 mL test water or 100 mL sediment with 400 mL test water). For single-concentration

tests, a minimum of 5 *replicate samples* of sediment (i.e., field replicates or separate samples from different grabs or cores taken at the same *sampling station*) must be taken at each discrete sampling station, and from each of one or more reference sampling stations (EC, 1992a, 1994, 1997a; USEPA, 1994a, 2000). These sediment samples must be tested for their toxicity to *H. azteca* as a single *replicate* (i.e., using only one test vessel per replicate sample; see Section 5.1).[34] For multi-concentration tests (e.g., spiked-sediment tests and water-only tests; see Sections 6.2 and 7.3, respectively) a minimum of 5 replicate vessels (i.e., laboratory replicates) per *treatment* must be established.[35] In all tests, a minimum of 5 replicates (i.e., laboratory replicates) must be established for the control sediment. The sediment added to each vessel should be smoothed using a spatula or by tapping the vessel against the side of the hand. Test water (see Section 3.4) is then poured slowly down the side of the beaker or jar. To minimize the disruption of sediment as test water is added, a disc made of Teflon™, polyethylene, or nylon sheeting, cut to fit the inside diameter of the test vessel, may be placed on the sediment surface before water is added[36] (EC, 1992a). A volume of test water (or, depending on the test, a test solution) should be added to the test vessel at this time such that the sediment to water ratio is 1:1.75 (e.g., 175 mL of water for 100 mL of sediment) or 1:4 (e.g., 400 mL of water for 100 mL of sediment); this can be judged using a mark inscribed at the required total volume (e.g., 275 mL or 500 mL) on the vessel's side. A somewhat lesser volume (e.g., 125 to 145 mL) of test water may be added initially, to provide room for any additional water added when test organisms are introduced. For certain special applications (e.g., site-specific or research), the investigator(s) might wish to use a larger sediment:water ratio (i.e., 1:67) in order to maintain a static exposure. The use of this method might be particularly applicable under special circumstances where more stable overlying water chemistry is desirable, such as for long-term testing (e.g., 28- or 42-days), including bioaccumulation or reproductive tests.[37]

[34] In the first edition of EPS 1/RM/33, a minimum of five replicate vessels (i.e., laboratory replicates) were required for single-concentration testing of field-collected sediment (see Section 5.1).

[35] USEPA (1994a, 2000) indicates a minimum requirement of four replicates per treatment, and recommends eight replicates per treatment for sediment-toxicity tests.

[36] A length of nylon monofilament line (or non-toxic equivalent) could be attached to the disc, to enable its removal once the test water is added. Alternatively, the disc could be cut from a polyethylene bag in a keyhole configuration that provides a circle with an attached portion for removal. The disc should be rinsed with test wate
if reused to prepare replicates of a treatment.
A separate disc should be used for each treatment.

[37] Larger ratios of sediment:water might be used for special purposes (e.g., site-specific or research applications). The "cone method" with a 1:67 sediment:water ratio, developed at the National Water Research Institute (NWRI) to overcome challenges created by particularly unstable overlying water (Borgmann and Norwood, 1999), has been used extensively in 4-week survival and growth tests (Borgmann *et al.,* 2001a, b, 2004; Norwood *et al.*, 2009). It has also been used in bioaccumulation tests (Borgmann and Norwood, 2002; Borgmann *et al*., 2004; Nowierski *et al.*, 2005), and 8- to 10-week sediment reproduction tests carried out under static conditions (Borgmann, 2002; Bartlett *et al*., 2004; Bartlett and Brown, 2011). The use of large sediment:water ratios (i.e., 1:67 or 15 mL sediment and 1000 mL overlying water) in an Imhoff settling cone (i.e., a 1-L funnel-shaped polycarbonate or glass container usually used for measuring the volume of suspended solids) negates the need for water renewal due to the large volume of overlying water, and a reasonable sediment depth (~2.3 cm) is maintained due to the shape of the cone. This method is described in detail elsewhere (Borgmann and Norwood, 1999; Borgmann, 2002; Borgmann *et al.*, 2005a).

For the static-renewal exposure, the water overlying the sediment in each test vessel (i.e., reference, test, and laboratory-control sediments) must be renewed a minimum of 3 times weekly, on non-consecutive days throughout the test (i.e., 6 times during the 14-day test), at a rate of 2 volume additions in 24 hours.[38] A replacement rate greater than 2 volume additions in 24 hours should be avoided to prevent unnecessary flushing and depletion of any contaminants that might leach from the sediment into the overlying water. The overlying water can be replaced manually or with the use of suitable apparatus enabling the timed and periodic automatic renewal of the overlying water in each test vessel at the appropriate rate (see Section 3.1). If an automated system is used, it should be calibrated before the test is started to verify its performance; flow rates through any 2 test vessels should not differ by more than 10% at any time during the test (USEPA, 1994a, 2000; ASTM, 1995a, 2010). If overlying water is renewed by siphoning, care should be taken to prevent disturbance of the sediment or accidental loss of amphipods emerged from the sediment during this procedure. No more than 90% of the water should be siphoned and replaced, and the end of the siphon must not contact the sediment.

The overlying water in each test vessel should be aerated overnight before the test organisms are introduced, as well as throughout the test (see Section 4.3). Each beaker or jar should be kept covered (watchglass or plastic lid) during the pretest and test periods, to minimize evaporation and to reduce the possibility of contamination. Any overlying water lost by evaporation should be replaced on Day 7 of the test (or more frequently, if desired or necessary) by the gentle addition of temperature-adjusted test water poured down the side of the test vessel. A mark (e.g., 275 mL) inscribed on the side of the test vessel can be used to judge this.

Test organisms used to begin the test are those that are 2- to 9-days old, and ranging in age by ≤3 days (≤2 days is recommended) on the day that the test is initiated (see Section 2.3). In many instances, the culture water and the water used as overlying water in the test will be the same, although this is not necessarily so. The objectives of a particular test might require the use of another water source (e.g., that from a particular site under investigation) as test water. If water other than culture water is used as the test water, *acclimation* of test organisms to this water is not required (USEPA, 1994a, 2000), although it might be advisable to do so in order to minimize any stress on the animals caused by different water quality characteristics. If test organisms are to be acclimated, a useful procedure is to hold them for 2 h in a 50:50 mixture of culture water:test water, then for 2 h in a 25:75 mixture of culture water:test water, followed by a final 2 h in 100% test water before their introduction to test vessels (Ingersoll and Nelson, 1990). Another useful procedure is to siphon off 20 to 30% of the culture water every 2 to 3 hours and replacing it with the test water, ensuring that the temperature remains constant over this acclimation period. This should be done on the day before the test starts.

If toxicity tests are intended using samples of estuarine sediment, it is recommended that the test organisms be acclimated gradually to estuarine water with a salinity similar to that of the *pore water* of the test sediments, before the start of the test. Alternatively, additional controls could be included in the study, using control sediment with a pore water salinity similar to that of the test sediments.

[38] This renewal rate is in keeping with that recommended by the USEPA (1994a, 2000).

On Day 0, 10 amphipods should be assigned randomly to each test vessel. These organisms should be handled as little and as carefully as possible (see Section 2.3.9) during their transfer to the test vessels. Amphipods must be placed below the air/water interface in the overlying water. Test organisms may be pipetted directly from a culture vessel into the overlying water (Ankley *et al.*, 1993a). Alternatively, 10 amphipods may be counted into a transfer vessel (e.g., 30-mL plastic cup) filled with test water at the test temperature, and then recounted before their transfer below the surface of the overlying water (Ingersoll and Nelson, 1990; USEPA, 1994a, 2000). The latter procedure is particularly useful, since it permits the organisms to be counted twice before they are introduced to the test vessel. Following the addition of test organisms, the volume of water overlying the sediment should be increased as necessary until the mark inscribed on the vessel's side (e.g., 275 mL) is reached.

4.2 Test Conditions

- This is a whole sediment toxicity test, during which the overlying water is normally not renewed except for the periodic addition of test water to replace that lost from evaporation (i.e., static test). If, however, the test water overlying any reference sediment is fouled (i.e., by rising ammonia levels) and/or deteriorates (due to drifting pH and/or a drop in DO) and the objectives of the test are to exclude these confounding effects from the overall toxic response, then the test solution must be renewed a minimum of 3 times weekly on non-consecutive days, at a rate of 2 volume additions in 24 hours (i.e., static-renewal test).

- Test duration is normally 14 days.[39]

- The test must be conducted at a daily mean temperature (overlying water) of 23 ± 1°C. Additionally, the instantaneous temperature must be 23 ± 3°C (USEPA, 1994a, 2000; ASTM, 1995a, 2010).

- The test vessel must be glass, and a 300-mL high-form glass beaker or glass jar with an inner diameter of ≥7 cm is recommended.

- Control and test sediments must be present as a uniform layer with a volume of overlying water such that the sediment to water ratio is the standard 1:1.75 (i.e., 100 mL sediment layer with a 175-mL volume of overlying water) or optionally 1:4 (e.g., 100 mL of sediment with 400 mL of overlying water) if a greater volume of overlying water is required for chemical analyses. A minimum sediment depth of 2 cm is required.

- Test vessels should be covered.[40] The overlying water in each vessel should be aerated continuously at a minimal rate (see Section 4.3). Organisms in each test vessel must be fed either three times per week (on non-consecutive days) or daily throughout the test (see Section 4.4).

- Test vessels are to be illuminated with a daily photoperiod of 16-h light and 8-h dark, using overhead, full-spectrum lights (fluorescent or

[39] See footnote 101.

[40] For tests where the overlying water is renewed, test vessels should be covered to minimize loss of volatiles from the sediment or to reduce the risk of contamination.

equivalent). Light intensity adjacent to the surface of the overlying water should be 500 to 1000 lux.

- For a valid test, the mean survival rate for amphipods in control sediment must be ≥80% at the end of the test. Additionally, the minimum dry weight for the replicate control groups at test end (Day 14) must average ≥0.10 mg per individual amphipod.[41]

4.3 Dissolved Oxygen and Aeration

H. azteca can tolerate hypoxic conditions (Section 1.4). The dissolved oxygen concentration in the overlying water in all test vessels including the controls should be maintained between 40% and 100% saturation. The overlying water in each test vessel should be aerated continuously on the night preceding the start of the test (see Section 4.1), as well as during the test. Compressed air, previously filtered so as to be free of oil, should be dispensed to each test vessel through airline tubing and a disposable plastic or glass tube (e.g., capillary tubing or a pipette having an Eppendorf tip) with a small aperture (e.g., 0.5-mm ID). The tip of the air delivery tube should be suspended approximately 3 cm above the surface of the sediment layer. Air flow to each test vessel must be gentle and regulated (e.g., 2 to 3 bubbles/s), and must not disturb the sediment surface (Zumwalt *et al.*, 1994). The air flow to each test vessel should be checked routinely (e.g., daily) throughout the test, and adjustments made if necessary to maintain a gentle rate of aeration. Any aeration during testing must be reported (Section 9).

4.4 Food and Feeding

Organisms in each test vessel must be fed either once daily, or three times weekly (on non-consecutive days) throughout the test. Since dry weight of amphipods is a primary endpoint for the test, an identical food ration must be added to each test vessel on each feeding occasion. The ration provided must be adequate to enable acceptable survival and growth of *H. azteca* during the test period (see Section 4.2), but must not be excessive.[42]

Throughout the test, *H. azteca* are fed one of three food options. The food options include:

i) an aqueous mixture of yeast, cereal grass media, and trout chow (YCT) (see Appendix H);

ii) finely ground commercial fish food flakes (e.g., Nutrafin®, Tetrafin®, TetraMin®, or Zeigler® Aquatox® Feed); or

iii) a 1:1 combination of YCT and finely ground commercial fish food flakes.[43]

[41] Based on a review of dry weight attained by control groups in 14-day tests with *H. azteca*, under static or static-renewal conditions defined in this report, Milani *et al.* (1996) concluded that a criterion for test validity of ≥0.10 mg per individual control organism would normally be attainable yet discriminatory for this species, and recommended this for inclusion as a test criterion using either option.

[42] Feeding during the test is essential to enable adequate (≥80%) survival and acceptable growth of test organisms (Ankley *et al.*, 1993a, 1994; Milani *et al.*, 1996). The addition of excess or different types of food is to be avoided since it might alter the bioavailability of contaminants in the sediment and/or promote the growth of fungi or bacteria on the sediment surface (USEPA, 1994a, 2000).

[43] These food types and rations have proven suitable for *H. azteca* under the defined test conditions (Borgmann *et al.*, 1989; NWRI, 1992; Ankley *et al.*, 1993a, 1994; USEPA 1994a, 2000; ASTM 1995a, 2010; Milani *et al.*, 1996; Hockett *et al*., 2011; P. Jackman, ALET, Environment Canada, Moncton, NB, personal communication, 2012). Other food types and rations, including single ration diets of rabbit chow (Ingersoll and Nelson, 1990; ASTM, 1991a, 1993), or multiple ration diets such as algae plus alfalfa plus fish

If daily feeding is chosen, a ration of 2.7 mg (dry weight) of food or equivalent (i.e., 2.7 mg of ground fish food flakes; or 2.7 mg of YCT; or 1.35 mg of YCT in combination with 1.35 mg fish flakes for the 1:1 mixture of YCT and fish food flakes) must be added daily to each test vessel on Day 0, as well as once per day thereafter until the day the test ends. If the option of feeding 3 times per week is chosen, a ration of 6.3 mg food, dry weight or equivalent (i.e., 6.3 mg ground fish food flakes; or 6.3 mg of YCT; or 3.15 mg of YCT in combination with 3.15 mg fish flakes for the 1:1 mixture of YCT and fish flakes) must be added 3 times per week (starting on Day 0) to each test vessel on non-consecutive days (e.g., on Mondays, Wednesdays, and Fridays) until the day the test ends. Test organisms are not fed on the last day (i.e., Day 14) of the test. Either ration results in the same overall rate of feeding; i.e., 18.9 mg dry food weekly, per test vessel. Daily feeding is preferable to "even out" the available food supply, although feeding 3 times per week might be a preferred choice to minimize weekend labour requirements (Milani *et al.*, 1996).[44]

Detailed records of the food type and ration added to each test vessel should be made on each feeding occasion. Observations of the appearance of the sediment surface in each test vessel (i.e., any evidence of a fungal or bacterial growth) should also be recorded at this time.

4.5 Observations and Measurements During the Test

Depending on the objectives, it might be worthwhile to regularly check each test vessel (preferably daily), to observe and record the number of amphipods seen swimming in the overlying water, floating on the water surface, or lying or grazing on the surface of the sediment.[45] Any animals seen floating on the water surface should be gently pushed down into the water using a glass rod or pipette.

The temperature of the overlying water must be measured at the beginning of the test, and thereafter at least three times per week on non-consecutive days (e.g., Mondays, Wednesdays, and Fridays) until test completion. These measurements must be made in at least one test vessel representing each *treatment*; and more frequent (i.e., daily) measurements are recommended.

food flakes, have been shown previously to enable adequate (≥80%) survival and acceptable growth of control animals using the conditions and procedures specified for this test. However, the use of a food type or ration other than that specified here is not recommended, since such differences could alter the bioavailability of contaminants and reduce the standardization of the test. Commercial fish food flakes should be finely ground (i.e., with a mortar and pestle) and passed through a 500–700 µm screen to ensure the flakes are ground finely enough for *Hyalella* to ingest, as well as uniform in size. The food may be prepared as an aqueous slurry (e.g., 2.7 mg mixed with 1.5 mL of water, or 1.35 mg mixed with 0.75 mL of YCT used to inoculate each test vessel each day), or sprinkled over the surface of the test vessels. If the food remains on the water's surface, it is not available to the *Hyalella*; therefore, care must be taken to ensure the food sinks to the bottom of each test vessel. See Appendix H for preparing YCT.

[44] Results for 14-day side-by-side comparisons using either feeding regime showed that survival and growth (dry weight at test end) of *H. azteca* did not differ significantly, regardless of whether the static or static-renewal options were used (Milani *et al.*, 1996; MESI, 2010).

[45] Records of numbers of animals emerged from the sediment might prove useful in assessing avoidance responses. However, since *H. azteca* is an epibenthic amphipod that frequently emerges from clean sediment, such observations are not necessarily worthwhile and are not required as part of this test method.

Additionally, it is recommended that the temperature of any water bath used, and/or of the air in a temperature-controlled room or chamber used for the test, be recorded continuously.

For at least one test vessel representing each treatment, the concentration of dissolved oxygen in the overlying water must be measured at the beginning of the test, and thereafter at least three times per week on non-consecutive days (e.g., Mondays, Wednesdays, and Fridays) until test completion. More frequent (e.g., daily; ASTM, 1995a, 2010) measurements might be advisable and would be warranted for sediments having a high oxygen demand. A probe and calibrated dissolved oxygen (DO) meter is recommended for these measurements. The probe must be inspected carefully after each reading to ensure that organisms have not adhered to it, and must be rinsed in deionized or distilled water between samples to minimize cross-contamination. If beakers or jars are aerated during the test (Section 4.3), the position of the tip of the pipette in each test vessel and the rate of aeration should be checked frequently and routinely, and adjustments made as necessary.

For each reference sediment, ammonia concentrations and pH in the overlying water must be measured at the beginning of the test, and thereafter at least three times per week on non-consecutive days (e.g., Mondays, Wednesdays, and Fridays) until test completion. More frequent (e.g., daily) measurements might be advisable and would be warranted for reference sediments producing high ammonia levels and/or rapid pH change in the overlying water.[46] For all other treatments (i.e., at least one test vessel representing each treatment or replicate sample, including control sediment), ammonia and pH in the overlying water must be measured at the beginning and end of the test. For each measurement of ammonia (see APHA *et al.,* 2005 for guidance), the concentration of un-ionized ammonia should be calculated based on the concurrent measurements of pH and temperature for the overlying water (Trussell, 1972; USEPA, 1985c, 1999; EC, 2008). The following formulae can be used to calculate the un-ionized ammonia concentration:

> Un-ionized ammonia (mg/L) =
> total ammonia (mg/L) x
> (1 / (1 + 10 pKa –pH))
> (Emerson *et al*., 1975)
>
> where:
> pH is that measured in the overlying water
> pKa (the acid dissociation constant of NH_4^+) = 0.09018 + 2729.92/T
> T = Temperature in Kelvin

For static-renewal exposures, water quality measurements should be conducted at the beginning and end of each renewal period, in both the fresh and the used overlying water just before it is changed, or just after it has been changed.

Conductivity (and salinity if appropriate) in the overlying water must be measured at the beginning and end of the test for at least one test vessel representing each treatment. Additionally, hardness and/or alkalinity

[46] The requirement to measure ammonia and pH in samples of reference sediment is related to the need to monitor changes in the overlying water quality. The static-renewal exposure must be initiated if the quality of the water overlying any reference sediment deteriorates due to high levels of ammonia and/or shifting pH, and the objectives of the test are to assess a toxic effect without the confounding effects of deteriorating water quality. The same rationale applies to more frequent monitoring of DO levels in the overlying water.

concentrations in the overlying water should be measured at the beginning and end of the test in at least one test vessel representing each treatment (USEPA, 1994a, 2000).

Conductivity and pH (and salinity if appropriate) may be measured using probes and calibrated meters. Ammonia may be measured using an ion-specific electrode and by extracting an aliquot of the overlying water for this analysis. As with DO measurements, any probe inserted in a test vessel must be inspected carefully immediately after each reading, and rinsed in deionized water between samples. For measurements of hardness, alkalinity, and ammonia requiring sample aliquots, samples of overlying water should be taken from extra replicates set up for monitoring purposes, or directly from the test vessel just before the addition of test organisms, and upon completion of the test.[47] No more than 10% of the volume of the overlying water in a test vessel (i.e., ≤17.5 mL for 175 mLs of overlying water or ≤40 mL for 400 mLs of overlying water) should be removed for this purpose. A pipette should be used carefully to remove water from a depth of about 1 to 2 cm above the sediment surface. The pipette should be checked to make certain that no amphipods are removed during the collection of these water samples.

The water quality measurements determined at the beginning and end of a test for each treatment are useful as they provide an indication of the influence of the sediment on overlying water quality during the test. If, for any treatment, a marked change (e.g., >50%; USEPA, 1994a, 2000) in one or more of these water quality variables is found between the initial and final measurements, a check on the conditions and procedures used in the test (e.g., static vs static-renewal and the frequency of overlying water renewal, if applicable) is recommended, together with a careful consideration of the physicochemical characteristics of the sediment used in the test.[48]

4.6 Ending a Test

The test is terminated after 14 days. Just before sieving the contents of a test vessel, all live and apparently dead amphipods in the water column or on the surface of the sediment should be pipetted from the test vessel. Individuals which are completely inactive but not obviously dead (e.g., not decomposing) should be held in test water within a petri dish or other suitable container, and examined closely at this time using a low-power microscope or hand-held magnifying glass. These individuals should be prodded gently with a sharp point to confirm that they show no sign of life (such as a pleopod twitch), and are then to be counted as dead.

Numbers of dead and surviving amphipods recovered by pipetting should be recorded and dead animals discarded. All live animals should be placed in a numbered weighing boat or similarly small holding receptacle containing sufficient test water for rinsing and holding the amphipods briefly until all of the survivors sieved from the sediment in the test vessel are added to and rinsed in it.

A consistent amount of time should be taken to sieve the contents of each test vessel and examine this closely for recovery of live or dead organisms. To ensure that the procedure

[47] For the 1:1.75 sediment to water ratio, it might be necessary to pool water samples from individual replicates, or to set up extra replicates to yield adequate volumes for these measurements. Alternatively, the 1:4 sediment to water ratio can be used to allow for more overlying water for chemical measurements.

[48] If an automated water-renewal system is being used (see Sections 3.1 and 4.1), its operation should be monitored daily. Any observations of water flow problems, or overflows in test vessels due to clogged drain screens, should result in immediate cleaning or other required maintenance.

used to recover amphipods is adequate, it is recommended that the laboratory personnel responsible for sieving the contents of test vessels previously demonstrate that they are able to retrieve an average of at least 90% of similar-sized *H. azteca* from sediment.[49]

The following technique, taken from USEPA (1994a, 2000), is recommended for sieving the contents of each test vessel. Other techniques or mesh sizes may also be used provided that they have been demonstrated in preliminary trials to allow the retrieval of test organisms.[50]

1. Pour approximately 50% of the overlying water through a #50 (300 μm) U.S. Standard mesh sieve.

2. Swirl the remaining water to suspend the upper ~1 cm of sediment. Pour the suspended slurry through the #50 mesh sieve. Using test water, wash the contents recovered on the sieve into a white tray or pan for inspection.

3. Using test water and a wash bottle, rinse the coarser sediment remaining in the test vessel through a #40 (425 μm) mesh sieve and wash the contents recovered on this sieve into a second tray or pan for inspection.

All live animals recovered from the overlying water or sediment in a single test vessel are counted and placed together in a numbered weighing boat or similarly small holding receptacle, and rinsed in test water to remove any sediment adhering to the carapace of the animal. The rinse should be brief, and no more than 10 minutes following introduction of the first amphipod. After rinsing, the group of surviving amphipods should be transferred to a clean, aluminum weighing boat that has been previously numbered, weighed, and held in a desiccator.[51]

Separate weighing boats, each containing the group of surviving amphipods recovered from each test vessel (replicate), are placed in an oven, and dried for 24 h at 60 ± 5 °C (NWRI, 1992). Upon removal from the oven, the boats are moved immediately to a desiccator. Following cooling, each boat should be individually and randomly removed from the desiccator, and weighed immediately[52] to the nearest 10 μg on a balance that measures accurately to this limit. Mean dry weight per amphipod that survived the test is calculated for each group[53] (see Section 4.7).

[49] USEPA (1994a, 2000) recommends a check on recovery capability used by Tomasovic *et al.* (1995), whereby test organisms are added to control sediment and their recovery determined after 1 h using the same technique as that employed for sieving the contents of test vessels at the end of the test.

[50] As an alternative to sieving, the sediment can be placed in one corner of a shallow, translucent tray on a light table. The tray is tilted, and the sediment washed downhill with a wash bottle and test water, to expose and count the amphipods (U. Borgmann, Fisheries and Oceans Canada, National Water Research Institute, Burlington, ON, personal communication, 1994).

[51] It might be advisable to oven-dry the weigh boats for at least 48 h to achieve a constant weight, since wax deposits associated with the weigh boats could otherwise provide weighing errors (G. Ankley, USEPA, Duluth, MN, personal communication, 1994).

[52] The dried amphipods can take up water vapour readily, so weighing should be rapid and the time standardized among boats. At the same time, care must be taken because rapid movement and static charge could cause dried specimens to be lost from the weighing boat.

[53] The body length of individual amphipods surviving at the end of the test has been used as an alternative measurement of growth in sediment toxicity tests with *H. azteca* (see Appendix E.9). USEPA (1994a, 2000) endorses determinations of either dry weight or body length as endpoints representing growth. Herein, dry weight is the recommended indicator of growth. Measurement of body length offers some additional advantage over dry-weight measurements, in that specimens can be preserved for subsequent analyses

During the series of dry-weight determinations for the groups of amphipods from a test, the first boat weighed should be returned to the desiccator and weighed again at the end of all weighings, as a check on gain of water by the boats in the desiccator to be weighed subsequently. The change should not be >5%; if it is, redrying of all boats for ≥2 h and reweighing might be carried out. A few weighing boats should be tared, dried, and weighed without amphipods, and results should conform to the laboratory's quality control standards.

4.7 Test Endpoints and Calculations

The biological endpoints for this 14-day sediment toxicity test are survival and dry weight. Reduced survival and/or lesser weight at test end are assessed by comparison with replicate reference and/or control groups (see Sections 5.6 and 6.5). The most sensitive of the two effects is taken as the definitive indication of toxicity.[54]

At the end of the 14-day exposure, the number of amphipods alive and number dead are recorded for each replicate including the control groups. The following two endpoints must be calculated for each treatment:

i) the mean (± SD) *percentage* of amphipods that survived during the exposure.[55]

ii) the mean (± SD) dry weight per surviving amphipod, calculated from the total weight of the group of survivors.

Missing individuals are assumed to have died and disintegrated during the test, and are included in the tally of dead individuals for a replicate. The total dry weight of the group of survivors in a replicate is then measured.

The test is invalid if the average percent survival for amphipods held in the control sediment for 14 days is <80% at the end of the test. The test is also invalid if the average dry weight for the replicate control groups is <0.1 mg per individual amphipod surviving at the end of the test.

The two most common possibilities for a typical *Hyalella* test design involve:

i) Multiple sampling stations, in which responses at one or more test site sampling station(s) are compared with those at a reference site sampling

(USEPA, 1994a, 2000) and data derived from individuals can be used for nested ANOVA and for appraising sexual maturation (Kemble *et al.*, 1994). In the future measurements of length may be substituted in this test for dry-weight measurements provided that studies demonstrate conclusively that length is as, or more, sensitive an indicator of growth. Results by Becker *et al.* (1995) provide supporting evidence in this regard. In Environment Canada's *survival-and-growth* test using fathead minnows (EC, 1992c), growth is based on mean dry weight alone, and length is not used as a criterion of effect due to evidence that increased body depth and weight of healthy individuals is not adequately reflected in gains in body length during the test. A similar phenomenon could occur during the present 14-day *survival-and-growth* test using *H. azteca*.

[54] An alternate measure of toxicity, which combines lethality and final weight, is *biomass*. To calculate this endpoint, the total dry weight of the surviving *Hyalella* is divided by the initial number of organisms (normally, 10). Currently, this endpoint is used in the fathead minnow larval test in both Canada (EC, 2011a) and the United States (USEPA, 2002), but biomass is not routinely applied to sediment toxicity tests.

[55] Calculation of the mean percentage assumes that each replicate started with the same number of test organisms. If this is not the case (i.e., a different number of organisms were used in each test vessel), then the average of the proportion (i.e., a weighted average) should be calculated.

station,[56] with other test sampling stations, or with the control sediment (i.e., single-concentration test). Hypothesis testing is frequently used in the statistical assessment and the common outcome is that a response at a sampling station is either "different" or "not different" from another sampling station.

ii) Multiple concentrations of a substance(s) or material(s) of interest, achieved by spiking a sediment, by mixing a test sediment (or similar particulate material) with clean sediment, or by testing multiple concentrations in a water-only toxicity test (see Section 7.0). The required *endpoints* for a multi-concentration test are *LC50* for survival and *ICp* for dry weight at the end of the test.

In a scenario where there are multiple sampling stations, an understanding of the strengths of various study designs is critical for the successful application of statistical tests. The study objectives should be clearly defined before data is collected, with an appreciation both for the power (ability to detect an effect) of the test design and the ease of interpretation of the results. In general, it is advantageous to limit the number of comparisons made, and this is typically done by choosing a test design and statistical tests that compare test sampling stations with a reference sampling station. Further gains in power can be made if a gradient can be assumed (i.e., samples collected in sequential order downstream or away from the point source; see Section P.4 in EC, 2005). In some cases, study objectives and test design may not have been given adequate attention before the collection of the data, and to compensate, investigators will perform a comparison among all possible sampling stations, maximizing the number of comparisons made. This is strongly discouraged, particularly when large numbers of sampling stations are involved, because undesirable effects on Type I and Type II error rates may occur; interpretation of results is often more difficult; and unwarranted focus may be given to particular comparisons after data has been collected.[57]

Detailed statistical guidance on hypothesis testing for both final dry weight (a quantitative measurement) and mortality (a *quantal* measurement) is provided in Section 5.6. The requirements for LC50 and ICp endpoints are outlined in Section 6.5.

4.8 Tests with Reference Toxicant(s)

The routine use of a *reference toxicant* or toxicants is necessary to assess, under standardized test conditions, the relative sensitivity of the culture(s) of *H. azteca*, and the precision and reliability of data produced by the laboratory. *Water-only* tests with one or more reference toxicants are most commonly used in conjunction with *survival-and-growth* tests which measure sediment toxicity to *H. azteca* (see Appendix F).

[56] Throughout this document, *reference site* is used to describe an area where there is clean sediment uninfluenced by the contaminant under study (i.e., reference sediment). A reference sediment must be collected for these comparisons, as described earlier in this Section. However, in the absence of a reference sediment, a control sediment may be substituted for any of the tests listed here.

[57] Zajdlik and Associates Inc. (2010) made this last point in the defense of the application of an overall test for significance: "All too often an observed difference catches the eye of the data analyst and a search begins to apply a statistical test to 'validate' the observed difference. This is an example of data snooping; conclusions made using this data analytic approach are suspect." This same flaw is apparent in poorly defined study designs, as described here.

Procedures for spiking sediment with chemical(s) and for conducting spiked sediment reference toxicity tests are available or being developed (Burton, 1991; Smith *et al.*, 1992b; Suedel *et al.*, 1993a, b; EC, 1995) and should see wider use in the future. A static, 96-h *water-only* reference toxicity test is recommended here for routine use with sediment toxicity tests using *H. azteca,* a practice followed by USEPA (1994a, 2000). This reference toxicity test may be supplemented or replaced with one or more spiked sediment tests with reference toxicant(s) after suitable procedures are standardized. Environment and Climate Change Canada's guidance document on using control sediment spiked with a reference toxicant should be consulted (EC, 1995).

Table 3 provides a checklist of conditions and procedures recommended and required for conducting static, 96-h *water-only* reference toxicity tests using *H. azteca*. The recommended test procedure, which is largely consistent with USEPA (1994a, 2000), uses 2- to 9-day old amphipods that range in age by ≤3 days to start the test. There are 10 individuals per test vessel, at least 5 test concentrations plus a control (i.e., using control/dilution water-only), and 1 or more replicates per treatment. Recommended test vessels are 300-mL, high-form glass beakers or glass jars with an inner diameter of ≥7 cm, and the recommended test volume is 200 mL solution/vessel. A substrate for the *Hyalella* must be added to each test vessel, and for a given test, the substrate used must be identical for each test solution and each replicate used in that test. Options for test substrate include: an ~3 cm^2 strip of medicinal gauze bandage, presoaked in culture water for 24 h; an ~3 cm^2 piece of Nitex® or plastic mesh (e.g. 500 µm),[58] or a thin layer (i.e., 1–2 mm deep; ~5 mL for the recommended 300 mL high-form glass vessels) of clean silica sand (see footnote 27). Solutions in test vessels are not aerated during the test and are normally covered to minimize contamination and losses due to evaporation. For the reference toxicity tests, *H. azteca* are fed either an aqueous mixture of yeast, cereal grass media, and trout chow (YCT; see Appendix H) or a 1:1 mixture of YCT and finely ground commercial fish food flakes (e.g., Nutrafin®, Tetrafin®, TetraMin® or Zeigler® Aquatox Feed). It is recommended that a ration equivalent to 0.9 mg (dry weight) of food (i.e., an ~0.5 mL inoculum of YCT, or 0.25 mL of YCT in combination with 0.45 mg fish flakes for the 1:1 mixture of YCT and fish flakes; see Sections 4.4 and 7.5.3) be added to each test vessel (including the controls) on Days 0 and 2 of the test.[59] Temperature and lighting conditions for this test procedure are the same as those described for definitive sediment toxicity tests (see Section 4.2 and Tables 2 and 3). Daily observations are made for numbers of dead or moribund amphipods in each test vessel. For the reference toxicant tests, temperature is measured daily for each treatment. Dissolved oxygen, conductivity and pH must be measured at the start in all treatments.

[58] Some chemicals (e.g., PAHs) can be readily adsorbed by nylon mesh, therefore the suspected (or known) toxicants must be taken into consideration when choosing a substrate for water-only tests.

[59] In the first edition of this test method document, there was only one food option for reference toxicity tests (i.e., 0.5 mL YCT). With the new food options included herein, there is a possibility that the overall performance of test organisms could be affected. Separate warning charts must be set up for each different food type used for reference toxicity testing, if the type of food being used impacts the performance of the test organisms.

Table 3 Checklist of Recommended and Required Conditions and Procedures for Conducting Water-only Reference Toxicity Tests Using *Hyalella azteca*

Test type	–	static 96-h water-only toxicity test
Reference toxicant	–	copper sulphate ($CuSO_4$), cadmium chloride ($CdCl_2$), potassium chloride (KCl), or sodium chloride (NaCl)
Frequency of test	–	perform within 14 days of test start, or concurrently with definitive sediment or water-only test(s); if *Hyalella* are imported, test organisms from this batch concurrently with definitive test(s)
Test solutions	–	control and at least five test concentrations
Solution replacement	–	none
Control/dilution water	–	culture water or other clean ground or surface water; reconstituted fresh water if a high degree of standardization is required; natural or reconstituted seawater with salinity ≤15‰ for tests with estuarine sediment; DO, 90 to 100% saturation when used in test
Amphipods	–	removed from known-age culture as <1- to 7-d olds and held in beaker for 2 d preceding test while fed; 2- to 9-d old, and ranging by ≤3 d at start of test; 10/test vessel
Substrate for amphipods	–	substrate required; must be identical for all test vessels; options include: a 3 × 3 cm strip of medicinal gauze bandage, presoaked in culture water for 24 h, a 3 × 3 cm piece of Nitex® or plastic mesh, or a 1–2 mm deep (i.e., ~5 ml) layer of clean silica sand
Test vessel	–	glass beaker or glass jar; recommend 300-mL high form ≥7 cm internal diameter; normally covered
Volume of test solution	–	200 mL
Number of replicates	–	one or more per concentration
Temperature	–	daily average, 23 ± 1°C; instantaneous, 23 ± 3°C
Lighting	–	overhead full-spectrum (fluorescent or equivalent); 500 to 1000 lux; 16-h light:8-h dark
Aeration	–	none unless dissolved oxygen in overlying water drops below 40% of saturation
Feeding	–	aqueous suspension of YCT; or a 1:1 mixture of YCT and ground commercial fish food flakes (e.g., Nutrafin®, Tetrafin®, or TetraMin®, or Zeigler® Aquatox Feed) equivalent of 0.9 mg (dry weight) of food added to each vessel on Days 0 and 2

Observations	–	daily, each vessel, for number of dead or moribund amphipods
Measurements of water quality	–	daily, each treatment, for temperature; start of test, each treatment, dissolved oxygen, conductivity and pH; alkalinity and hardness should be measured for each treatment at the start and end of the test
Endpoints	–	mean percent survival, each treatment; 96-h LC50
Test validity	–	results for reference toxicity test considered invalid if mean 96-h survival in control water <90%

Alkalinity and hardness should be measured for each treatment at the start and end of the test. The test endpoints are the mean percent survival in each treatment, and the 96-h LC50. Results for a reference toxicity test must be declared invalid if the mean survival in control water is <90% at the end of the test (Table 3).

Appropriate criteria for selecting suitable reference toxicants might include the following (EC, 1990, 1995):

- chemical readily available in pure form;
- stable (long) shelf life of chemical;
- can be interspersed evenly throughout clean substrate;
- good dose/response curve for test organism;
- stable in aqueous solution;
- minimal hazard posed to user;
- concentration easily analyzed with precision;
- known influence of water quality (e.g., pH, hardness) on toxicity of chemical to test organism; and
- known influence of physicochemical characteristics of sediment (e.g., particle size, organic carbon content) on toxicity of chemical to test organism.

Reagent-grade copper sulphate, cadmium chloride, potassium chloride, or sodium chloride are recommended for use with *H. azteca* as reference toxicants (USEPA, 1994a, 2000).

Reference toxicity tests using *H. azteca* and one or more of these chemicals must be within 14 days before or after the date that the toxicity test is initiated or by performing this test concurrently with the definitive one, using the laboratory's established cultures. The performance of any cultures recently established in the laboratory using new breeding stock should also be evaluated using reference toxicant(s) before these cultures are used to provide test organisms (see Sections 2.3.1 and 2.3.11).

If test organisms are imported to the testing laboratory, rather than selecting them from an in-house culture that is the recommended approach (see Section 2.2), a portion of the juveniles from each *batch* of imported

organisms must be tested for its tolerance to the reference toxicant(s). The reference toxicant test must be performed at the same time as the definitive test, following the procedures and conditions described herein in Section 4.8.

Pertinent reports by Environment and Climate Change Canada provide guidance on the selection, performance, and use of water-only (EC, 1990) or *spiked-sediment* (EC, 1995) reference toxicity tests. Laboratory personnel unfamiliar with such tests are advised to consult these reports before preparing for or conducting them.

It is the laboratory's responsibility to demonstrate its ability to obtain consistent, precise results with reference toxicant(s) before conducting definitive sediment tests with *H. azteca*. To meet this responsibility, the laboratory personnel should initially determine intralaboratory precision, expressed as coefficient of variation (CV), by performing five or more tests with the reference toxicant(s) using different groups of *H. azteca* from separate known age cultures (Section 2.3.10) reared in the laboratory. For these preliminary tests, the same reference toxicant(s), concentrations, type/source of test water, and test procedure (i.e., Table 3) should be used. Performance of the routine tests with reference toxicant(s) should continue to follow this same procedure. A series of test concentrations should be chosen,[60] based on preliminary tests, to provide partial mortalities in two or more concentrations and enable calculation of a 96-h LC50 with acceptably narrow confidence limits (see Section 6.5).

Once sufficient data are available (EC, 1990, 1995), LC50s for a particular reference toxicant must be plotted successively on a *warning chart*, and examined to determine whether the results are within ±2 SD of the mean LC50 values obtained in previous tests with *H. azteca* using the same reference toxicant and test procedure. A separate warning chart must be prepared and updated for each reference toxicant used. The warning chart should plot logarithm of concentration on the vertical axis against date of the test or test number on the horizontal axis. Each new LC50 for the reference toxicant must be compared with established limits of the chart; the LC50 is acceptable if it falls within the *warning limits*.

The logarithm of concentration (including LC50) must be used in all calculations of mean and standard deviation, and in all plotting procedures. This simply represents continued adherence to the assumption by which each LC50 was estimated based on logarithms of concentrations. The warning chart may be constructed by plotting the logarithmic values of the mean and ±2 SD on arithmetic paper, or by converting them to arithmetic values and plotting those on the logarithmic scale of semi-log paper. If it were demonstrated that the LC50s failed to fit a log-normal distribution, an arithmetic mean and SD might prove more suitable. The *geometric mean* LC50, together with its respective upper and lower warning limits (±2 SD), should be recalculated with each successive LC50 for the reference toxicant until the statistics stabilize (EC, 1990, 1995, 2005). If a particular LC50 fell outside the warning limits, the sensitivity of the test organisms and the performance and precision of the test would be suspect. Since this might occur 5% of the time due to chance alone, an outlying LC50 would not necessarily indicate abnormal sensitivity of the culture or

[60] See Appendix I for guidance in selecting an appropriate series of test concentrations. Each successive concentration chosen should be at least 50% of the previous concentration.

unsatisfactory precision of toxicity data. Rather, it would provide a warning that there might be a problem. A thorough check of the health of the culture (Section 2.3.11) together with all culturing and test conditions should be carried out. Depending on the findings, it might be necessary to repeat the reference toxicity test, to obtain new breeding stock, and/or to establish new *known age* cultures, before undertaking further sediment toxicity tests. Results that remained within the warning limits might not necessarily indicate that a laboratory was generating consistent results. Extremely variable data for a reference toxicant would produce wide warning limits; a new data point could be within the warning limits but still represent undesirable variation in test results. A coefficient of variation of no more than 30%, and preferably 20% or less, is suggested as a reasonable limit by Environment Canada (1990).

Section 5

Specific Procedures for Testing Field-Collected Sediment or Similar Particulate Material – 14-Day Survival and Growth

This section gives particular instructions for the collection, preparation, and testing samples of field-collected sediment or similar particulate material. These instructions are in addition to the procedures listed in Section 4. Toxicity tests with these samples must be conducted using the static or static-renewal method (i.e., depending on the quality and stability of the test water overlying reference sediment, and the objectives of the test) described in Section 4. The daily-renewal test method may be used only under special circumstances described in Section 4 or as dictated by related regulatory guidelines or requirements.

Detailed guidance for the collection, handling, transport, storage, and analyses of field-collected sediment is given in ASTM (1991b, 1995b, 2008), EC (1994), and USEPA (2001) reports specific to these subjects. Environment Canada (1994) should be consulted and followed (in addition to the guidance provided here), when collecting samples of field-collected sediment and preparing them for toxicity tests with *H. azteca*.

5.1 Sample Collection

Environment Canada (1994) provides a useful summary of field-sampling design and appropriate techniques for sample collection. Field surveys of sediment toxicity using biological tests with *H. azteca* and/or other suitable, sediment-associated test organisms are frequently part of more comprehensive surveys. Such surveys could include a battery of toxicity tests to evaluate the toxicity of whole sediment, pore water, or elutriate, together with tests for bioaccumulation of contaminants, chemical analyses, biological surveys of epifaunal and/or infaunal organisms, and perhaps the compilation of geological and hydrographic data. Statistical correlation can be improved and costs reduced if the samples are taken concurrently for these tests, analyses, and data acquisitions.

Samples of sediment collected for assessment of an adverse effect on survival and growth of *H. azteca* might be routinely taken (e.g., quarterly, semi-annually, or annually) from a number of sites and/or sampling stations for *monitoring* and *compliance* purposes, or might be collected on one or more occasions during field surveys of sites for spatial (i.e., horizontal or vertical) or temporal definition of sediment quality. One or more sites must be sampled for reference (presumably clean) sediment during each field collection.[61]

The number of sampling stations to be sampled at a study site and the number of replicate samples per station will be specific to each study. This will involve, in most cases, a compromise between logistical and practical constraints (e.g., time and cost) and statistical considerations. Environment

[61] A reference sediment is that collected near the site(s) of concern. Ideally, it possesses geochemical characteristics similar to those of the test sediment but without anthropogenic contaminants. It is not unusual for nearby reference sites to have some degree of contamination due to anthropogenic chemicals. In some instances, reference sediment might be toxic due to naturally occurring physical, chemical, or biological properties (Burton, 1991).

Canada (1994) should be consulted for guidance with respect to the sampling design.

A minimum of five replicate samples (i.e., field replicates or separate samples from different grabs or cores taken at the same station) of sediment must be taken at each discrete sampling station. Sample collection must also include a minimum of five replicate samples (i.e., field replicates) from each of one or more reference sampling stations (EC, 1992a, 1994, 1997a; USEPA, 1994a, 2000). The objective of collecting replicate samples[62] at each sampling station is to allow for statistical comparisons within and among different stations (EC, 2005). Accordingly, each of these "true replicate" samples of sediment must be tested for its toxicity to *H. azteca* as a single replicate (i.e., using only one test vessel per replicate sample).[63] The use of *power analysis* (see Section 5.6.4) with endpoint data obtained in previous tests at the same or similar sites will assist in determining if more than five replicate samples need to be tested. For certain other purposes (e.g., preliminary study or extensive surveys of the spatial distribution of toxicity), the survey design might include only one replicate sample from each station. The latter approach precludes any determination of mean toxicity at a given sampling location (station), but allows a statistical comparison of toxicity of each sample with the control, and also if desired, a comparison among the test samples (stations), using appropriate statistical tests (see Section 5.6.1). Sites for collecting *reference sediment* should be sought where the geochemical properties of the sediment are similar to sediment characteristics encountered at the test sites. Close matching of sediment grain size or organic content might not be necessary for this test, since *H. azteca* can tolerate uncontaminated sediments differing in these properties without changes in survival or growth (Section 1.4). Matching of organic carbon content might not be warranted in cases where pollution (e.g., from pulp mills or sewage) is responsible for the high organic content of test sediments. Preliminary surveys to assess the toxicity and geochemical properties of sediment within the region(s) of concern and at neighbouring sites are useful for selecting appropriate sites at which to collect reference sediment.

Samples of municipal or industrial sludge (e.g., sewage sludge, dewatered mine tailings, sludge from an industrial clarifier or settling pond) might be collected for assessment of their adverse effect on survival and growth of *H. azteca*, and for geochemical and contaminant analyses. Other particulate wastes (e.g., drilling mud residue) might also be taken for toxicity and chemical evaluation.

Procedures used for sample collection (i.e., core, grab, or dredge) will depend on the study objectives and the nature of the sediment or other particulate material being collected. The types of sediment collection devices and their advantages and disadvantages have been summarized by Environment Canada (1994), and further details are provided elsewhere (de Groot and

[62] *Replicate sample(s)* are field-replicated samples of sediment collected from the same sampling station, to provide an estimate of the sampling error or to improve the precision of estimation. A single sediment sample from a sampling station is treated as one replicate. Additional samples collected at the same sampling station are considered to be additional replicate samples and must be treated identically but stored in separate sample containers (i.e., not composited).

[63] Data shows that the testing of replicate vessels (i.e., laboratory replicates), when nested inside replicate samples (i.e., field replicates) have minimal impact on the power of an analysis (i.e., the ability to detect an effect). This trend extends to most data sets (i.e., the "higher up" you go in a nested design, the more impact your replicates will have). To balance scientific, practical and cost considerations replicate vessels are no longer required for testing samples of field-collected sediment in single-concentration tests.

Zschuppe, 1981; Baudo *et al.*, 1990; Burton, 1992; Sly and Christie, 1992; ASTM, 1995b, 2008).

A benthic grab or core rather than a dredge should be used for sampling sediment, to minimize disruption of the sample. Sediment to be evaluated for toxicity and chemistry should be collected from one or more depths that represent the layer(s) of concern (e.g., a surficial 2-cm layer, or a deeper layer if there are concerns about historical deposition of contaminants).

Care must be taken to minimize loss of fines during sample collection. If the sample is obtained using a grab sampler, hand corers should be used to collect a sample from the surficial 2 cm, or desired layer, of the test sediment retrieved using this sampling device. This can be achieved if the grab can be opened from the top to expose the surface of the undisturbed sediment. The sample should be transferred to a clean sample container.

Before commencing a sampling program, the required volume of sediment per sample should be calculated (EC, 1994). This calculation should take into account the quantity of sediment required to prepare a single replicate for sediment toxicity tests, as well as that required for particle size characterization, percent organic matter, percent moisture, and specific chemical analyses. A volume of at least 1 L of sediment per sample is normally required (EC, 1994), although this will depend on the study objectives/design and the nature of the chemical analyses to be performed. To obtain the required sample volume, it might be necessary to combine subsamples retrieved using the sampling device. Guidance provided in Environment Canada (1994) for compositing subsamples in the field should be followed.

The same collection procedure should be used at all field sites and stations sampled. Environment Canada (1994) should be consulted for further guidance on appropriate devices and procedures for sample collection.

5.2 Sample Labelling, Transport, Storage, and Analyses

Containers for transport and storage of samples of field-collected sediment or similar particulate material must be made of non-toxic material. The choice of container for transporting and storing samples depends on both sample volume and the potential end uses of the sample. Environment Canada (1994) should be consulted for guidance in selecting suitable containers. The containers must either be new or thoroughly cleaned, and rinsed with test water or other clean water (e.g., distilled or deionized water) before use.

Each sample container should be filled completely, to exclude air. Immediately after filling, each sample container must be sealed, and labelled or coded. Labelling and accompanying records made at this time must include at least a code or description that identifies sample type (e.g., grab, core), source, precise location (i.e., water body, latitude, longitude, depth), replicate number, and date of collection; and should include the name and signature of sampler(s). Persons collecting samples of sediment should also keep records describing details of:

- the nature, appearance, and volume of each sample;
- the sampling procedure and apparatus;
- any procedure used to composite or subsample grabs or cores in the field;
- the number of replicate samples taken at each sampling station;

- the sampling schedule;
- the types and numbers of containers used for transporting samples;
- any field measurements (e.g., temperature, salinity, pH, dissolved oxygen) of the overlying water or sediment at the collection site; and
- procedures and conditions for cooling and transporting the samples.

Upon collection, warm (>7°C) samples should be cooled to between 1 and 7°C with regular ice or frozen gel packs, and kept cool (4 ± 3°C) in darkness throughout the period of transport. As necessary, gel packs, regular ice, or other means of refrigeration should be used to assure that the temperature of the sample(s) remains within 1 to 7°C during transit.

The date of receipt of the sample(s) at the laboratory must be recorded. Sample temperature upon receipt at the laboratory should also be measured and recorded. Samples to be stored for future use must be held in airtight containers and in darkness at 4 ± 2°C (EC, 1994, 1997a). Any air "headspace" in the storage container should be purged with nitrogen gas, before capping tightly (EC, 1994). Samples must not freeze or partially freeze during transport or storage, and must not be allowed to dry (EC, 1992a, 1994, 1997a). It is recommended that samples of sediment or similar particulate material be tested as soon as possible after collection. The sediment toxicity test should begin within two weeks of sampling, and preferably within one week; the test must start no later than six weeks after sample collection, unless it is known that the sediment contaminants are stable (i.e., will not change appreciably).[64]

Ideally, sediment characteristics that are unstable (e.g., pH, oxidation-reduction potential) or changed by conditions of transit and storage (e.g., temperature) should be measured in the field to help characterize the sample. In the laboratory, each sample of field-collected sediment should be thoroughly mixed (Section 5.3), and representative subsamples taken for physicochemical characterization. Each sample (including each field replicate and all samples of control and reference sediment) must be characterized by analyzing subsamples of whole sediment for at least the particle size distribution (percentage of coarse-grained sand, medium-grained sand, fine-grained sand, silt, and clay) and total organic carbon content.[65] In addition, the measurement of pore water and/or whole sediment pH and ammonia (total and un-ionized concentrations; see Section 4.5), and percent water content for each sample is recommended.[66] Other

[64] The toxicity and geochemistry of contaminated sediments from Hamilton Harbour were reported to change with storage for longer than one week, although the data supporting that statement were not provided (Brouwer *et al*., 1990). A study by Othoudt *et al.* (1991) found that the toxicity of samples of freshwater sediment did not differ significantly when stored at 4°C for periods of 7 to 112 days. Burton (1991) and USEPA (1994a, 2000) report studies by various researchers showing in some instances that the toxicity of sediment held at 4°C was unchanged after several months' storage, and in other cases that changes were noted within days to weeks. A recommendation for testing within two weeks conforms with the advice in other sediment toxicity tests by Environment Canada (1992a, 1997a). A maximum permissible storage time of 6 weeks has been recommended by Environment Canada (1994) for sediments intended for toxicity tests, in view of practical difficulties for shorter times, including time required if initial chemical analyses are to be performed.

[65] Measurements of ammonia and pH in the overlying water are also required for each treatment at the beginning and end of the test (see Section 4.5).

[66] The sediment chemistry requirements here focus on variables that may impact an organism's health directly. For example, there has been some anecdotal evidence that sediment with a high nutrient content can result in increased size of test organisms, even though test conditions require feeding. Some experts have

analyses could include (USEPA, 1994a, 2000; APHA *et al.*, 1995, 2010): total inorganic carbon, total volatile solids, biochemical oxygen demand, chemical oxygen demand, cation exchange capacity, acid volatile sulphides, metals, synthetic organic compounds, oil and grease, petroleum hydrocarbons, and pore water analyses for various physicochemical characteristics. Unless indicated otherwise, identical chemical, physical, and toxicological analyses should be performed with subsamples representative of each replicate sample of field-collected sediment (including reference sediment) taken for a particular survey of sediment quality, together with one or more subsamples of control sediment.

5.3 Preparing Samples for Testing

Field-collected sediment or similar particulate material should normally not be prepared for testing by sieving with water, as this would remove contaminants present in the pore water or loosely sorbed to particulate material (EC, 1994). Large debris or large indigenous macro-organisms should normally be removed using forceps or a gloved hand.

The presence of indigenous macro-organisms in samples can reduce the growth of *H. azteca* in freshwater sediment toxicity tests, and can confound the interpretation of test results (Reynoldson *et al.*, 1994). If a field-collected sediment contains a large number of indigenous macro-organisms that cannot be removed using forceps, the sample may be press-sieved (not washed) through one or more suitably sized mesh screens. For those sediments containing small macro-organisms that due to sediment characteristics cannot be removed by press-sieving, the sample(s) may be rinsed through a fine-mesh sieve (e.g., 0.25 to 0.5 mm; Day *et al.*, 1995b) using any liquid that has separated from the sample during its transport and/or storage. This liquid must be remixed within the sieved sample (Section 4.1).

Sieving could alter the concentration or bioavailability of contaminants in the sediment, or alter its nutrient content and/or particle size (EC, 1994; Day *et al.*, 1995b). If sediments are sieved, therefore, it is recommended that the physicochemical properties of the sediment (e.g., pore water metals, particle size distribution) be documented before and after sieving. Comparative toxicity tests using sieved and unsieved sediment might, in some cases, also be necessary or appropriate to discern the effect of sieving on sample toxicity. Unless research or special study objectives dictate otherwise, each sample of field-collected test material should be homogenized in the laboratory before use (EC, 1994; USEPA, 1994a, 2000).[67] Mixing can affect the concentration and bioavailability of contaminants in the sediment, and sample homogenization might not be desirable for all purposes.

suggested that the pH and ammonia of overlying water (see Section 4.5) are adequate for the assessment of direct affects on *Hyalella*, an epibenthic species. As a result, pore water and/or whole sediment measurements of pH and ammonia are not required, but recommended. Note that this minimum chemistry does not take into account contaminant interactions. It is the responsibility of the investigator to measure physicochemical variables that can act as toxicity modifying factors (e.g., EC, 2010). Since each contaminant situation is unique, it is beyond the scope of this method to mandate the measurement of contaminants or toxicity modifying factors.

[67] One of the reasons for routinely homogenizing samples is to mix into the sediment any pore water that rises to the surface during sample shipment and storage. Homogenization is also necessary to redistribute the sample constituents that have compacted and layered according to particle size during transport and storage.

To achieve a homogeneous sample, either mix it in its transfer/storage container, or transfer it to a clean mixing container. The sample may be stirred using a non-toxic device (e.g., stainless steel spoon or spatula) until its texture and colour are homogeneous (EC, 1992a). Alternatively, a mechanical method (USEPA, 1994a, 2000; EC, 1994) may be used to homogenize the sample. For each sample included in a test, mixing conditions including duration and temperature must be as similar as possible. If there is concern about the effectiveness of sample mixing, subsamples of the sediment should be taken after mixing and analyzed separately to determine homogeneity.

Immediately following sample mixing, subsamples of test material required for the toxicity test and for physicochemical analyses must be removed and placed in labelled test vessels (Section 4.1), and in the labelled containers required for storage of samples for subsequent physicochemical analyses. Any remaining portions of the homogenized sample that might be required for additional toxicity tests using *H. azteca* or other test organisms should also be transferred to labelled containers at this time. All subsamples to be stored should be held in sealed containers with no air space, and must be stored in darkness at 4 ± 2°C (Section 5.2) until used or analyzed. Just before it is analyzed or used in the toxicity test, each subsample must be thoroughly remixed to ensure that it is homogeneous.

5.4 Test Water

For tests with field-collected sediment or similar particulate material, the water introduced to test vessels (i.e., overlying water) may be from the same source as that used for culturing *H. azteca* (see Sections 2.3.4 and 3.4). Alternatively, this water may be from a separate supply of natural fresh or estuarine water, or reconstituted water. For certain applications, the experimental design might require or endorse the use of fresh or estuarine water taken from the reference site near where test sediments were collected. Use of uncontaminated site water, or uncontaminated water adjusted to the hardness of site water, is frequently a good choice due to the modifying influence of waters with different hardness values on the toxicity of metals or organic contaminants in sediment. When site water is used as overlying water, a second set of controls must be prepared using a supply (source) of laboratory water shown previously by the testing laboratory to routinely enable valid test results in a 14-day test for survival and growth of *Hyalella* (see Section 3.4). Section 2.3.4 provides pertinent guidance on the preparation and analysis of water to be used as overlying water in the test.

5.5 Test Observations and Measurements

A qualitative description of each field-collected test material should be made when the test is being set up. This might include observations of sample colour, texture, and homogeneity; and the presence of plants, animals, and tracks or burrows of animals (EC, 1992a). Any changes in the appearance of the test material and in the overlying water observed during the test or upon its termination should be noted and reported.

Measurements of the quality of the overlying water (e.g., pH, conductivity, salinity temperature, hardness, alkalinity, ammonia, dissolved oxygen content) in test vessels should be made during or at the beginning and end of the test, as described in Section 4.5. Depending on the test objectives and experimental design (i.e., sediment to water

ratio), separate test vessels might also be set up at the beginning of the test (Section 4.1) to monitor whole sediment and/or pore water chemistry (USEPA, 1994a, 2000). These would be destructively sampled during and at the end of the test. Test organisms might or might not be added to these extra test vessels, depending on study objectives. Measurements of chemical concentrations in the sediment or pore water within these vessels may be made by siphoning most of the overlying water without disturbing the surface of the sediment, then removing aliquots of the sediment for the appropriate analyses (see Section 5.2). If pore water were to be analyzed, centrifugation without filtration would be the recommended sampling procedure (EC, 1994; USEPA, 1994a, 2000). Environment Canada (1994) should be consulted for guidance on the recommended procedure for extracting pore water, and its treatment and storage before analyses.

Depending on the study objectives and the nature of the test sediments (e.g., rich in organics), measurements of pore water pH and ammonia concentrations might be made as the test progresses, using test vessels dedicated for this purpose (EC, 1994; USEPA, 1994a, 2000). Other sediment characteristics (e.g., concentrations of metals, hydrogen sulphide, total volatile solids, Eh) might be monitored in the same test vessels. If it were desired to monitor these variables, at least one vessel should be set up for each treatment, and destructively sampled for this purpose.

5.6 Test Endpoints and Calculations

The common theme for interpreting tests with samples of sediment is a comparison of the biological effects in one or more test sediments with the effects found in a reference sediment. A reference sediment should be used for comparative purposes whenever possible or appropriate, because this provides a site-specific evaluation of toxicity (USEPA, 1994a, 2000). Sometimes the reference sediment might be unsuitable for comparison because of toxicity or atypical physicochemical characteristics. In such cases, it would be necessary to compare the test sediments with the control sediment. Control sediment(s) results will assist in distinguishing contaminant effects from non-contaminant effects caused by such things as particle size and organic carbon content. Regardless of whether the reference sediment or control sediment is used for the statistical comparisons, the results from control sediment must be used to judge the validity and acceptability of the test (Section 4.7).

The two required measurement endpoints in the *Hyalella* test (see Section 4.7) are mortality (a *quantal* measurement) and dry weight, as an indication of growth (a *quantitative* measurement), at the end of the test. Because of the different nature of the measurements involved, different statistical approaches are needed, and these approaches are further refined to reflect the objectives of the experiment. This section will provide statistical guidance for data from single-concentration tests (i.e., sediments from multiple sampling stations tested at full strength only) considering three common test scenarios:

i) comparison of one test sampling station and one reference sampling station (see Scenario 1, Figure 2);

ii) comparison of several *ordered*[68] sampling stations with one reference

[68] The term *ordered* in this context indicates that there would be an expected gradient along the sampling stations, such as a series of sampling stations located progressively further from a point source. In this context, there is no measurement associated with the ordered stations (e.g., no measured distance from the

sampling station (see Scenario 2, Figure 2); and

iii) comparison of several *unordered* sampling stations with one reference sampling station (see Scenario 3, Figure 2).

Less-common test design scenarios are considered in Section 5.6.3. Only summary guidance for analysing the dry weight endpoint is provided here (Section 5.6.2), as more extensive statistical guidance is available elsewhere (EC, 2005). Although not required, statistical analysis (hypothesis testing) of test data in a site comparison context is recommended. If the collected reference sediment proves to be unsuitable (e.g., after physicochemical analysis in the lab), investigators may wish to use control sediment for comparisons with test sampling stations. Note that this will result in a test design that mixes replicate vessels (control sediment) with replicate samples (test sampling stations), and following the guidance herein, obliges the investigator to treat replicate vessels as equivalent to replicate samples. While this is not appropriate statistically, it will need to be considered acceptable, given the lack of reasonable alternates. If inferences drawn from the analysis are deemed to be of high impact (e.g., clean-up criteria), a statistician should be consulted.

Multi-concentration tests might be conducted with sediment, sludge, or similar particulate material, where measured amounts of the test sample could be mixed with measured quantities of natural or formulated control sediment (see Sections 3.5 and 6.2). Procedures for mixing different samples of sediment are not yet standardized or proven (see Section 6.2), and caution is advised due to possible non-linear responses and changes in bioavailability or sorption characteristics (Nelson *et al.*, 1994). A minimum of five test vessels are recommended to provide replicates (i.e., laboratory replicates) in each concentration, to determine sample homogeneity and test precision. Statistical analyses to determine endpoints for multi-concentration tests are described in Section 6.5.

5.6.1 Analysis of Mortality Data

Historically, there has been limited guidance on the analysis of quantal data in a test design that examines multiple sampling stations (EC, 2005). Environment and Climate Change Canada has recently improved its guidance for analysis in this scenario (Zajdlik & Associates Inc., 2010), and a summary is presented here (see Figure 3). In general, the preferred method of analysis is *logistic regression* followed by *contrast analysis*. If logistic regression is not available,[69] more widely available tests that are easily implemented are suggested.

source), so the independent variable (sampling station) is characterized as *ordinal*. *Ordered* is equivalent to the use of "gradient expected" (EC, 2005). "*Unordered*" assumes that there is no such gradient of responses, and more generally is characterized as "categorical." These test design distinctions (e.g., ordinal, categorical) are determined during the experimental design phase (*a priori*), not after the data has been collected.

[69] As of May 2012, the statistical program most often used by Canadian laboratories (CETIS) does not yet have the capability to perform logistic regression or subsequent contrast analyses.

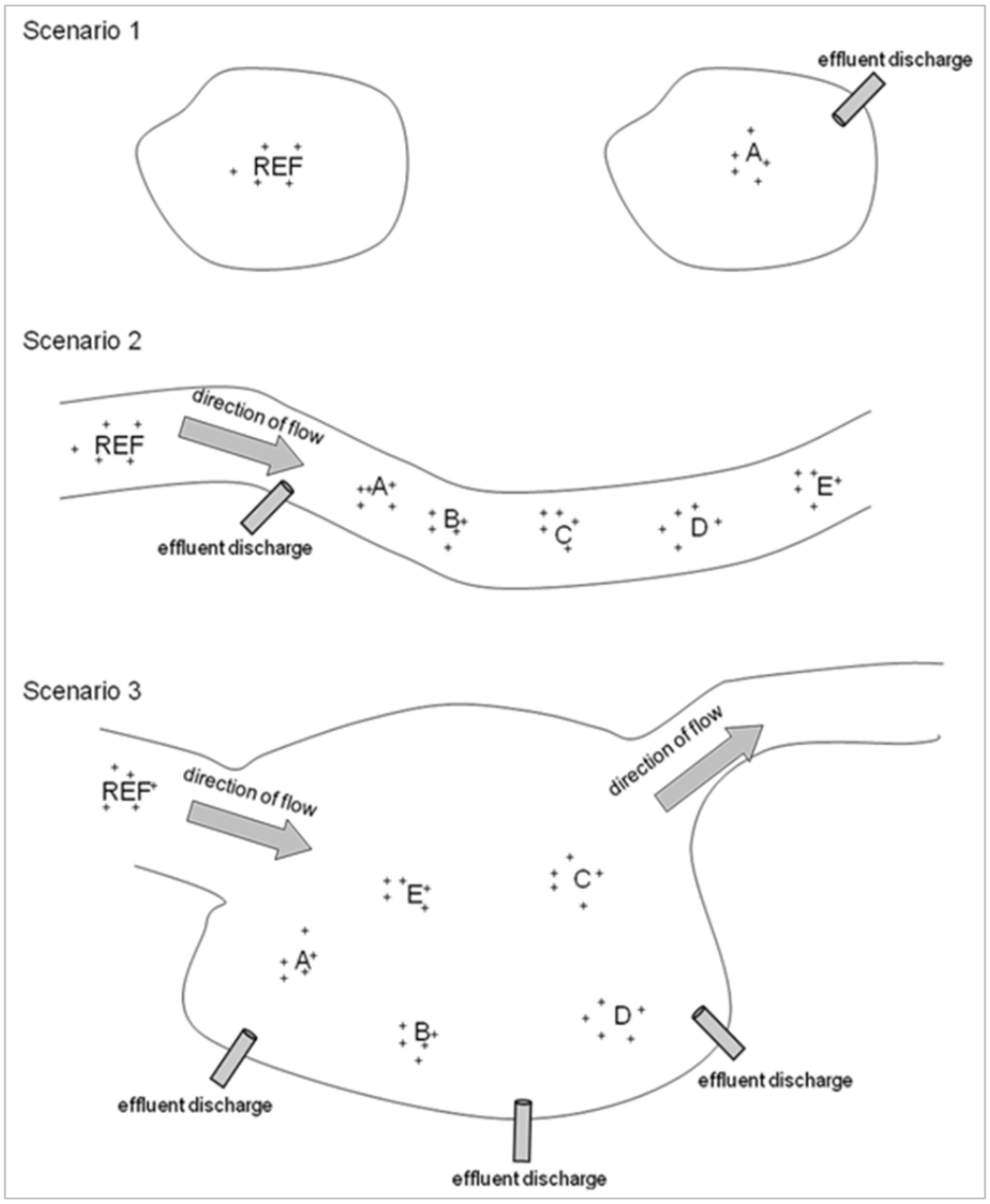


Figure 2 Common Test Designs Used in Sediment Field Evaluation

Schematic showing examples of common test designs used in sediment field site evaluation. Scenario 1 shows a test design with a reference sampling station (REF) and a test sampling station (A). Scenario 2 shows one reference sampling station (REF) with five test sampling stations (A–E) that are ordered. Scenario 3 shows one reference sampling station (REF) with five test sampling stations (A–E) that are unordered. In all three scenarios, each sampling station contains five replicate samples (field replicates) (+).

With only one test sampling station and one reference sampling station, and assuming replicates[70] are taken, the preferred test is logistic regression. If logistic regression is not used, the replicates can be combined (i.e., summed) and *Fisher's exact* (next preferred) or a *chi-squared* test[71] (least preferred) are acceptable. If, as a result of a special application (e.g., preliminary study), replicates are not collected (see Section 5.1), the preferred test is *Fisher's exact*, and an acceptable alternate would be a *chi-squared* test.

To compare several ordered test sampling stations with a reference sampling station, the preferred test is logistic regression followed by contrast tests for comparing individual sampling stations.[72] If the treatment-response is linear (i.e., even "spacing" between the variables), a sequential testing procedure using contrast analysis is recommended (Zajdlik & Associates Inc., 2010). If logistic regression is not used, the next preferred option is the *Cochran-Armitage* trend test, and a last acceptable alternative would be the *chi-squared* test.[71] The *post-hoc* alternatives to the sequential contrast procedure for a monotonic treatment-response[73] are *Shirley's* test (next preferred) or the *Jonckheere-Terpstra* test (least preferred). There may be a situation where, although a gradient response was expected, this was in fact not observed (i.e., non-monotonic treatment-response). This may occur, for example, if downstream from a point-source, the nutrient status of sediment changes, or the bioavailability of the contaminant does not remain consistent. If this is observed, then the appropriate test would be *Fisher's exact* with a *Bonferroni-Holm* adjustment (preferred) or the *Wilcoxon* rank sum test.[74]

In the third situation, there are multiple test sampling stations, but these are not ordered along an expected gradient, and comparisons are made with a reference sampling station.

The preferred test is logistic regression followed by paired contrast statements with a *Bonferroni-Holm* adjustment. If logistic regression is not available, a *Fisher-Freeman-Halton* would be the next preferred choice, or a *chi-squared* test[71] (least preferred) could also be used. An alternative post-hoc test would be paired *Fisher's exact* test with a *Bonferroni-Holm* adjustment.

5.6.1.1 Assessment of Models Used

If logistic regression is used, the suitability or adequacy of the model to explain the observed data is often assessed. The two tools for assessing adequacy of the model are:

i) test of significance of the explanatory variable, and

ii) test of model fit.

[70] Unless otherwise stated, all replicates are considered to be *field replicates*.

[71] If available, application of a continuity correction for the *chi-squared* test is recommended. If there is only one test sampling station and one reference sampling station (i.e., a 2 × 2 contingency table), then the appropriate correction is *Yates* continuity correction. For test designs with more than one test sampling station, more general continuity corrections are appropriate.

[72] The logistic regression is used here as an overall test of significance, and contrast analysis is the post hoc test.

[73] *Monotonic treatment-response* describes a relationship between the ordinal sampling stations and the biological response in which the direction of response does not reverse when examined along the direction of sampling stations. For example, if mortality gradually decreased (or stayed the same) at successive sampling stations downstream from a point source, the relationship would be described as a monotonic treatment-response.

[74] This test is also known as the *Mann-Whitney U* test and the *Mann-Whitney-Wilcoxon* test.

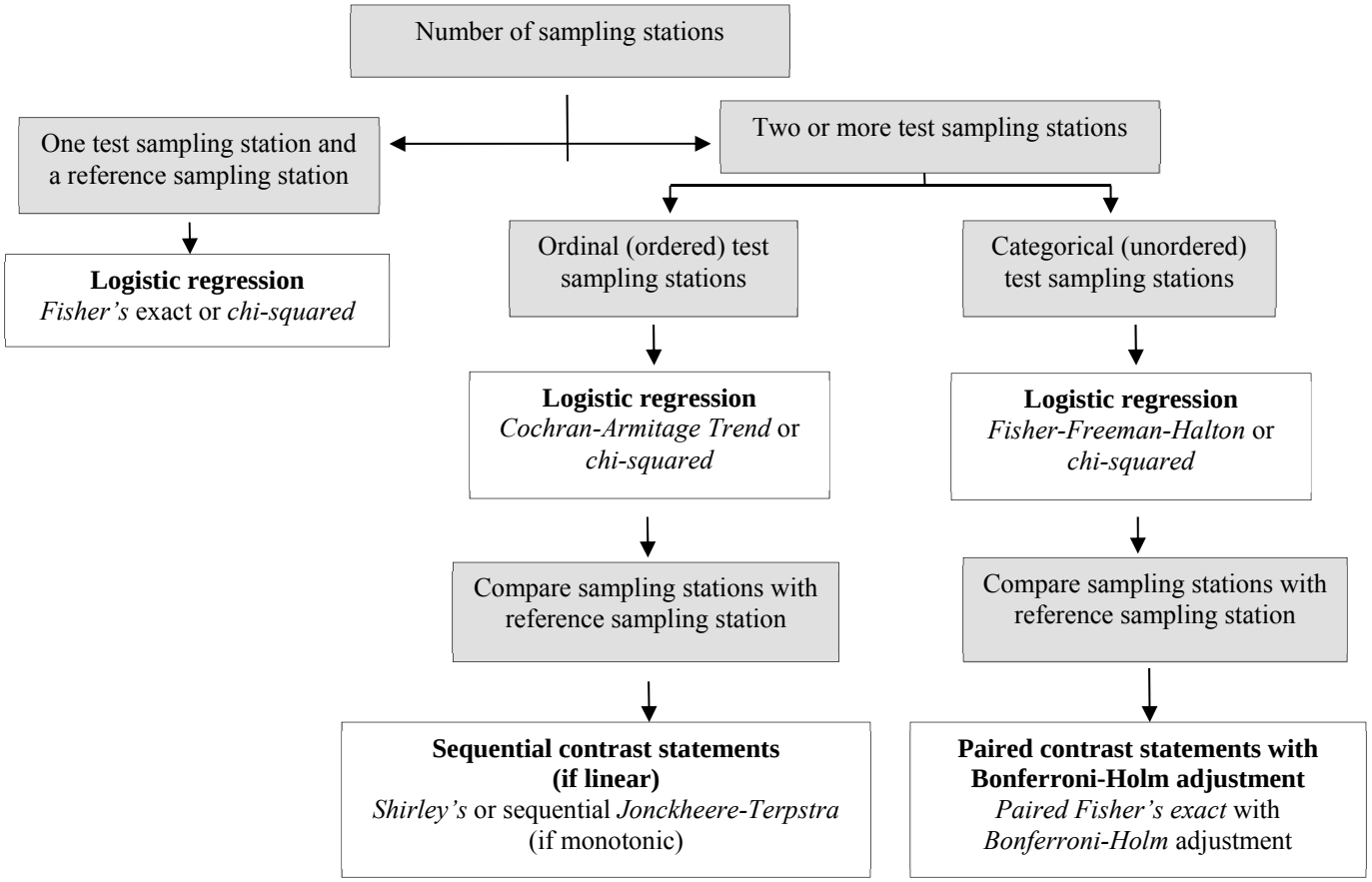


Figure 3 The Process for the Statistical Analysis of Three Common Test Designs in Sediment Field Testing

Test design features or experiment objectives (shaded boxes) that guide the decision path include number of sampling stations, nature of sampling stations (*ordinal* or categorical), and type of comparisons sought (here, only comparisons with a reference sampling station). Use of bold font indicates the preferred test. Among alternates listed (regular font), consult the main text for any preferred tests. Tests using contingency tables (e.g., *Fisher's exact*, *chi-squared*, *Fisher-Freeman-Halton*) normally require combining field replicates. Yates (or similar) continuity correction is used with *chi-squared* tests. All scenarios above assume replicate samples (field replicates) are used, and a single reference sampling station is used. If a reference sampling station is unavailable, control sediment may be substituted. If sampling stations are assumed to be *ordinal* during the test design phase, but subsequent analysis of test data do not support a *monotonic treatment-response*, sequential contrast statements, *Shirley's* test and sequential *Jonckheere-Terpstra* tests may not be appropriate; consult main text for alternate tests.

In the test designs described here, the test for significance of the explanatory variable can be described as testing whether differences between sampling stations can adequately explain the differences in mortality observed. Suitable tests for assessing this significance are the *Likelihood ratio* test (preferred) or *Wald* test. To examine model fit,[75] plots of residuals are visually assessed. Poor model fit may occur when the responses do not follow a binomial distribution[76] or when outliers are present.

Decision criteria and full interpretation for test of significance of the explanatory variable and test of model fit have not yet been established, and alternate options for assessing the adequacy of the model have not been explored. As such, use of these tools for Quality Assurance /Quality Control purposes is left to the knowledgeable user, or a statistician may be consulted.

Non-parametric tests (such as *Shirley's* or the *Jonckheere-Terpstra* test) are not subject to test of significance of the explanatory variable or tests of model fit. However, there are underlying assumptions common to all statistical tests, such as independence of responses. In properly designed and executed standardized tests, these assumptions are unlikely to be violated (Zajdlik & Associates Inc., 2010) and so are not discussed further.

5.6.2 Analysis of Dry Weight Data

Environment and Climate Change Canada has provided detailed statistical guidance on the analysis of quantitative measurements (EC 2005)[77] that can be readily applied to measurements of *Hyalella* growth (i.e., final dry weight) in a multiple sampling station scenario. If final dry weight at a single test sampling station is to be compared with final dry weight at a reference sampling station, a *t-test*[78] is normally the appropriate statistical test (Section 3.2 in EC 2005). In situations where more than one test sampling station (treatment) is under study, and the investigator wishes to compare multiple sampling stations with the reference, or compare sampling stations with each other, a variety of ANOVA and multiple comparison tests (and non-parametric equivalents) exist (Section 3.3 in EC, 2005). Choice of a specific test depends on:

i) the type of comparison that is sought (e.g., complete series of pairwise comparisons between all sampling stations, or compare the response from each sampling station only with that of the reference site);

ii) if a chemical and/or biological response gradient is expected; and

iii) if the assumptions of *normality* and *homoscedasticity* are met.

5.6.3 Variations in Design and Analysis

Less common design scenarios in sediment testing include:

[75] Model fit describes the ability of the model to accurately predict the response variable.

[76] Although it is clear that the response with *Hyalella* mortality is binomial, there is the possibility that other sources of variation (for example, replicates within a sampling station) would contribute variation to the model. There are formal tests for a binomial distribution (Zajdlik & Associates Inc., 2010), but these are not deemed critical for use at this point.

[77] Sections 3.2 and 3.3 in EC, 2005 provide guidance on the analysis of quantitative measurements for a single location and for multi-locations, respectively, and should be consulted for the analysis of growth data. Section 7.5 in EC, 2005 provides additional guidance on Multiple-comparison tests for hypothesis testing, and should be consulted for additional detail; however, the calculation of *NOEC*/*LOEC* is not recommended herein.

[78] The *t-test* assumes equal variance between groups; however, modification of the t-test that can accommodate unequal variance are also available (EC, 2005).

i) comparison of all sampling stations with each other (full pairwise comparison);[79]

ii) comparison of sampling stations that have been grouped into "near field" and "far field" categories;[80]

iii) collection of replicate vessels and replicate samples.

For a full pairwise comparison (e.g., comparison of all sampling stations with each other) of final dry weight data (quantitative), and assuming data meet the assumptions of normality and equality of variance, an ANOVA would first be conducted to test for overall differences. Post-hoc tests, such as *Fisher's Least Significant Difference* or *Tukey's* test, could then be used. Sections 3.3 and 7.5 in EC, 2005 provide further details, alternate tests, and non-parametric options, and the guidance therein should be followed.

For a full pairwise comparison (e.g., comparison of all sampling stations with each other) of mortality data (quantal), choice of overall test for significance is identical to those described for the multiple sampling stations scenarios in Section 5.6.1 (i.e., logistic regression is first choice). In the case where data are *ordered*, conflicts in interpretation of a multiple sampling station comparison may occur, and the investigator is encouraged to reconsider hypotheses testing (Zajdlik & Associates Inc., 2010). If data are *unordered*, then the multiple comparison procedure follows that for comparison with controls (preferred method is paired contrast statements, with an alternative of paired *Fisher's exact* with *Bonferroni-Holm* adjustment).

Guidance for comparing "near-field" and "far-field" sampling stations in the case of final dry weight data has not been developed expressly for this Biological Test Method. The reader is referred to other Environment and Climate Change Canada guidance, which provides details on other quantitative endpoints, given several different study designs (Chapters 4 and 8 in EC, 2011b). Note that terminology used to describe study areas in these documents may not correspond to terminology used in this Biological Test Method.

If the sampling design and experimental objectives allow the grouping of several sampling stations into "near-field" and "far-field" groups, then a partitioning of the c*hi-squared* test is recommended for mortality data (quantal), as this approach can increase the ability to detect a significant difference. Calculation details are listed elsewhere (Zajdlik & Associates Inc., 2010).

If both *replicate samples* (i.e., field replicates) and *replicate vessels* (i.e., laboratory replicates) have been tested, a statistician should be consulted for analysis options.

5.6.4 Power Analysis

An important factor to consider in the analysis of toxicity tests with sediment is the potential for declaring false "positives" (i.e., calling a clean site contaminated; Type I error) or false

[79] This would indicate a full pairwise comparison, which is broader in scope than the comparison of test sampling stations only with a reference sampling station, as outlined in Section 5.6.1.

[80] For the purposes of this document, near-field and far-field areas are defined in an effluent context (EC, 2010). Near-field areas are exposure areas outside the zone where the effluent is directly released into the environment (the effluent discharge zone), and have the highest effluent exposure. Far-field areas are exposure areas located further down the effluent gradient, and have a lower concentration of effluent than the near-field exposure areas. Conceptually, far-field exposure areas extend until reference conditions exist. Near-field and far-field areas are often compared with each other and may combine more than one sampling station. A reference area (i.e., reference site) may also be compared in this manner.

"negatives" (i.e., calling a contaminated site clean; Type II error). Scientists are usually cautious in choosing the level of significance (α) for tolerating false positive results (Type I error), and usually set it at p = 0.05 or 0.01. Commonly, scientists following a specified test design will never consider the relationship between power, variability, and effect size, leaving the Type II error completely unspecified. There are several factors that influence statistical power, including:

- variability of replicate samples representing the same treatment;
- α (i.e., the probability of making a Type I error);
- effect size (*ES*) (i.e., the magnitude of the true effect for which you are testing); and
- *n* (i.e., the number of samples or replicates used in a test, and in some cases, the allocation of those replicates[81]).

In research-based science, power analysis is most useful as part of a preliminary test design (Hoenig and Heisey, 2001; Lenth, 2007; Newman, 2008). Here, a preliminary experiment is run to determine the approximate standard deviation (variability), and to troubleshoot the execution of the experiment in general. Other factors in power analysis, such as effect size and number of replicates, can then be considered along with the standard deviation so that the final test design is optimized (e.g., number of replicates needed to detect a certain effect size is determined).

In the development of standardized test methods, the purpose of employing power analysis remains the optimization of test design (or at least estimating the power of the current test design).[82] However, instead of a single estimate for variability and effect size, there would typically be a much richer data set to consider. For example, test method experts could collect a number of estimates of variability, across different laboratories and different contaminant scenarios (Thursby *et al*., 1997; Van der Hoeven, 1998; Denton *et al*., 2011). Standardized tests are often used in *monitoring* or regulatory programs, which may specify the expected effect size (e.g., 25%) to be detected (Porebski and Osborne, 1998).

[81] If the experimental design requires the comparison of test sampling stations with the reference sampling station only (e.g., using *Dunnett's* test or *Williams'* test), optimal power for the final dry weight endpoint is achieved by allocating a higher number of replicate samples at the reference sampling station (Dunnett, 1955; Williams, 1972; OECD, 2006). As a general rule, the number of replicate samples at the reference sampling station (n_o) can be related to the number of test sampling stations (k) and the number of replicate samples at each test sampling station (n) using: $n_o = n\sqrt{k}$ for *Dunnett's* test (OECD, 2006). A modified version is recommended if *Williams'* test is used, where $\sqrt{k}$ is replaced with a range between $1.1\sqrt{k}$ and $1.4\sqrt{k}$ (Williams, 1972). With the current test method, each sampling station must have a minimum of 5 replicate samples. If the investigator was interested in increasing the number of replicate samples beyond the minimum, extra replicate samples should be allocated to the reference sampling station to maximize power and minimize Type II error. As an example using *Dunnett's* formula, consider an experiment with reference sampling station and 4 test sampling stations, and each test sampling station with 5 replicate samples. To maximize power, the optimal number of replicate samples at the reference sampling station would be $n_o = n\sqrt{k} = 5*\sqrt{4} = 10$ replicates.

[82] In 2010, the USEPA introduced a data analysis approach termed the test of significant toxicity approach (TST; USEPA 2010). The TST is a hypothesis testing approach based on bioequivalence, which is extensively used in pharmaceutical development and evaluation. We include it in discussions here because power analysis and the TST share some similar goals (e.g., a priori statement of Type I and Type II error) and because of the similar context (application of standardized testing).

The long-term goal of Environment and Climate Change Canada is to collect this data, and use it to assess the ability of the current test design to achieve a defensible Type II error rate. However, because field replicates replaced laboratory replicates in the second edition test method, adequate estimates of standard deviation are not yet available.[83] Based on a single assessment, that used the growth endpoint and a model that incorporated both field and laboratory replicates (Zajdlik & Associates Inc., 2010), a very preliminary assessment of power (estimated for five field replicates) can be determined. Based on this data, the power to detect a 30% inhibition in growth is estimated to be between 65% and 85%.[84]

The reproduction endpoint (see Section 8) is inherently more variable than growth or survival endpoints (ASTM, 2010) and for this reason, more replicates are required than the standardized 14-day exposures. The 42-day method requires a minimum of 8 replicates per treatment and 10 replicates are recommended. Power analysis completed with method validation data obtained to date from three laboratories (AquaTox Testing & Consulting Inc. (Guelph, ON), Environment and Climate Change Canada's Atlantic Laboratory for Environmental Testing (Moncton, NB, Canada) and Pacific and Yukon Laboratory for Environmental Testing (North Vancouver, BC, Canada) showed the ability of this test design to detect 40% effect with 90% power when 8 replicates were used per treatment. With ten replicates, a 30% effect was detected with 80% power based on average standard deviations for reproduction. Therefore, choosing the number of replicates will depend on the goals of individual studies. The power to detect smaller effects will also likely increase as laboratories gain experience with this method (Taylor *et al*., 2016).

[83] Variability between field replicates cannot be assessed from historical data, which used laboratory replicates.

[84] The range of power estimates reflects the range of variance estimates used. In this case, variance components from field replicates and laboratory replication were combined in a ratio. It is this variance ratio that was used to estimate power.

Section 6

Specific Procedures for Testing Spiked Sediment – 14-Day Survival and Growth

This section gives guidance and instructions for preparing and testing control or other sediment spiked experimentally with chemical(s), contaminated sediment, or complex waste mixtures. These recommendations and instructions are in addition to the procedures listed in Section 4. More detailed and appropriate guidance for spiking sediment with chemical(s) and conducting toxicity tests with chemical/ sediment mixtures is given in Environment Canada (1995). Further evaluation and standardization of procedures for spiking sediment (Section 6.2) might be required before sediment toxicity tests using *H. azteca* or other appropriate test organisms are applied to evaluate specific chemical/sediment mixtures for regulatory purposes.

The cause(s) of sediment toxicity and the interactive toxic effects of chemical(s), contaminated sediment, or particulate waste in association with otherwise clean sediment can be examined experimentally by spiking *clean* control sediment with these substances or materials. The spiking might be done with one or more chemicals, with another sediment (clean or uncontaminated), or with similar particulate material (e.g., dredged sludge mixed with sediment from an existing or prospective freshwater disposal site).

Toxicity tests using sediment spiked with a range of concentrations can estimate LC50s, and can determine concentrations causing *sublethal effects*. The influence of the physicochemical characteristics of natural or formulated sediment on chemical toxicity can also be determined with spiked-sediment toxicity tests. Reference toxicity tests can also be conducted using control sediment spiked with an appropriate chemical (see Section 4.8). Specific recommendations and instructions for performing spiked-sediment tests are provided in this section (additional useful guidance is given in USEPA, 1994a, 2000 and EC, 1995).

6.1 Sample Properties, Labelling, and Storage

Information should be obtained on the properties of the chemical, contaminated sediment, or particulate waste to be diluted experimentally with control or other sediment. For samples of contaminated sediment or similar particulate material, instructions on sample characterization (Section 5.2) should be followed. For individual chemicals, chemical substances (e.g., formulated products), or chemical mixtures, available information should be obtained on the concentration of major ingredients and impurities, water solubility, vapour pressure, chemical stability, dissociation constants, toxicity to humans and aquatic organisms, and biodegradability. Where aqueous solubility is in doubt or problematic, acceptable procedures previously used for preparing aqueous solutions of the chemical(s) should be obtained and reported, and chemical solubility in test water should be determined experimentally. Other available information such as structural formulae, nature and percentage of significant impurities, presence and amounts of additives, and n-octanol:water

partition coefficient should be obtained and recorded. [85]

Chemical(s) to be tested should be at least reagent grade, unless a test on a formulated commercial product or technical grade chemical(s) is required (USEPA, 1994a, 2000). Chemical containers must be sealed and coded or labelled upon receipt. Required information (chemical name, supplier, date received, person responsible for testing, etc.) should be indicated on the label and/or recorded on a separate datasheet dedicated to the sample, as appropriate. Storage conditions (e.g., temperature, protection from light) are frequently dictated by the nature of the chemical. Sample(s) of contaminated sediment or particulate waste to be evaluated in spiked-sediment assays should be collected, labelled, transported, and stored according to instructions herein (Sections 5.1 and 5.2).

6.2 Preparing Test Mixtures

Different procedures have been used by researchers to spike (dose) *clean* control sediment with chemical(s), or to dilute contaminated sediment or other particulate waste with control sediment, in preparation for sediment toxicity tests with the mixture (ASTM, 1991a, b, 1993, 1995a, b, 2008, 2010; Burton, 1991; USEPA, 1994a, 2000, 2001; Hoke *et al.*, 1995).

Procedures used to prepare spiked sediment (including substance or material addition and mixing, equilibration time and conditions) are varied. Mixing technique and time, as well as the period of aging after mixing, can affect the toxicity of the mixture (USEPA, 1994a, 2000). Accordingly, a standardized methodology for preparing spiked sediment cannot be recommended at this time; however, some of the approaches used previously or thought to be reasonable for preparing spiked sediment for toxicity tests with *H. azteca* are given here.

Reports from Environment Canada (1994, 1995), USEPA (2000, 2001), and ASTM (2008) provide more detailed instructions and recommendations for spiking and homogenizing sediment, and should be consulted for further guidance. Researchers intending to pursue toxicity tests using one or more laboratory-prepared mixtures should proceed cautiously, and should be well aware of potential problems due to non-homogeneity of the mixture(s) and the associated changes in bioavailability/sorption characteristics and non-linear toxic responses that might result (Nelson *et al.*, 1994).

The method to be used for experimentally spiking sediment is contingent on the study objectives and the nature of the test substance or material to be mixed with control or other sediment. In many instances, a chemical/ sediment mixture is prepared by making up a *stock solution* of the chemical and then mixing one or more measured volumes into control sediment (Swartz *et al.*, 1985b, 1988; ASTM, 1991a, 1993). Chemical concentrations in sediment are frequently calculated and expressed as µg/g or mg/kg dry weight (Swartz *et al.*, 1985b, 1988), although concentrations based on wet weight might be more useful for relating results to sediment toxicity (Burton, 1991). Depending on the nature of the test substance or material and test objectives, concentrations might also be normalized to sediment organic carbon content (e.g., for evaluating the toxicity of non-polar organic compounds) or to acid volatile sulphides (e.g., for assessing metal

[85] Knowledge of the properties of the chemical will assist in determining any special precautions and requirements necessary while handling and testing (testing in a well-ventilated facility, need for solvent, etc.). Information regarding chemical solubility and stability in water will also be useful in interpreting results.

toxicity) (Di Toro *et al.*, 1990, 1991; ASTM, 1991a, 1993; USEPA, 1994a, 2000).

The preferred solvent for preparing stock solutions is test water (see Sections 2.3.4 and 3.4); use of a solvent other than water should be avoided unless it is absolutely necessary. For organic compounds or other chemicals that do not dissolve readily in test water, a water-miscible organic solvent may be used to help disperse the compound in water (Borgmann *et al.*, 1990; ASTM, 1991a, 1993; USEPA, 1994a, 2000). Triethylene glycol has been recommended because of its low toxicity to aquatic organisms, low volatility, and high ability to dissolve many organic chemicals (ASTM, 1991a, 1993). Other solvents such as dimethylsulphoxide, methanol, ethanol, or acetone may be used to prepare stock solutions of organic chemicals, although they might contribute to sample toxicity, alter sediment properties, or be lost from the test material due to their volatility. Surfactants should not be used (ASTM, 1991a, 1993).

If an organic solvent is used, the test must be conducted using both a clean sediment control (i.e., no solvent and no test substance) and a sediment control containing solvent. For this purpose, a *solvent control sediment* must be prepared containing the concentration of solubilizing agent that is present in the highest concentration of the test chemical in sediment. Solvent from the same batch used to make the stock solution must be used (ASTM, 1991a, 1993; USEPA, 1994a, 2000).

Solvents should be used sparingly, as they might contribute to the toxicity of the prepared test sediment. The maximum concentration of solvent in the sediment should be at a concentration that does not affect the survival or growth of *H. azteca* during the test. If this information is unknown, a preliminary *solvent only* test, using various concentrations of solvent in control sediment, should be conducted to determine the threshold-effect concentration of the particular solvent being considered for use in the definitive test. To reduce the possibility of solvent-related artifacts, the spiking process should include a step that allows the solvent to evaporate before addition of sediment and water (USEPA, 2001; ISO, 2011). Using a shell coating technique, the chemical (dissolved in a solvent) can be coated on the walls of a container or onto silica sand. The solvent is slowly evaporated before the addition of the wet sediment. The wet sediment then sorbs the chemical from the dry surface (USEPA, 2001).

Measured volumes of a stock solution containing test chemical(s) should be mixed with control (or other) sediment in a manner resulting in a homogeneous distribution of the chemical(s) throughout the sediment. Spiking of both wet and dry sediment is common; however, the wet-spiking procedure is recommended since dry-spiking that can lead to losses of the test chemical and/or changes in its physicochemical characteristics, and that of the sediment. The USEPA (2001) recommends several methods for spiking sediment. Mixing may be by hand (e.g., using a clean spatula or glass rod), or by using the sediment rolling technique (e.g., using a mixing device; Ditsworth *et al.*, 1990). Alternatively, the sediment suspension technique (Cairns *et al*., 1984; Stemmer *et al*., 1990; Landrum and Faust, 1991) or the slurry spiking technique (Birge *et al*., 1987) are recommended (USEPA, 2001; ISO, 2011). Other methods of mixing might prove to be acceptable provided that the chemical is shown to be evenly distributed in the sediment. Mixing conditions, including solution:sediment ratio, mixing and holding time, and mixing and holding temperature, must be standardized for each treatment included in a test. Time for mixing a spiked

sediment should be adequate to ensure homogeneous distribution of the chemical, and may be from minutes up to 24 h. During mixing, temperature should be kept low to minimize changes in the mixture's physicochemical characteristics and microbial activity. Analyses of subsamples of the mixture are advisable to determine the degree of mixing and homogeneity (Ditsworth *et al*., 1990; USEPA, 1994a, 2000).

For some studies, it might be necessary to prepare only one concentration of a particular mixture of control (or other) sediment and chemical(s), or a mixture of only one concentration of contaminated sediment or particulate waste in control or other sediment. For instance, a single-concentration test might be conducted to determine whether a specific concentration of chemical or particulate waste in clean sediment is toxic to *H. azteca*. Such an application could be used for research or regulatory purposes (see Section 5.6 for guidance on statistical approaches for single-concentration tests).

A multi-concentration test, using a range of concentrations of chemical added to a control or other sediment under standardized conditions, should be used to determine the endpoint (e.g., LC50, ICp; see Section 6.5) for chemical/sediment mixtures. A multi-concentration test using control sediment spiked with a specific particulate waste might also be appropriate. For such purposes, at least 7 test concentrations plus a control must be prepared, and the preparation and use of more (i.e., at least 10 concentrations plus the control sediments) is recommended to improve the likelihood of attaining each endpoint sought. An appropriate geometric dilution series may be used, in which each successive concentration of chemical or particulate waste in sediment is at least 50% of the previous (e.g., 320, 160, 80, 40, 20, 10, 5, 2.5, 1.25, 0.63 mg/kg).[86] Test concentrations may also be selected from other appropriate logarithmic dilution series (see Appendix I). To select a suitable range of concentrations, a preliminary or range-finding test covering a broader range of test concentrations may be conducted.

Tests intended to evaluate the toxicity of mixtures of test substance(s) or material(s) in control sediment for federal registration or other regulatory purposes must be set up using a minimum of five replicates for each test concentration and each control sediment to be included in the assay. Since the objective for a multi-concentration test is to determine both LC50 (mortality data) and ICp (dry weight data), a test using 10 concentrations plus control(s) is recommended. The number of replicates per treatment could be reduced or eliminated altogether for range-finding tests and, depending on the expected variance among test vessels within a treatment, could also be reduced or eliminated for non-regulatory screening assays or research studies.

A period of equilibration is required to allow time for the test chemical concentration to stabilize in the sediment pore water spaces. The duration of this equilibration period is highly dependent on the nature of the test chemical and the sediment type (e.g., equilibration period in a sandy sediment or for an inorganic salt would be less than in a sediment with a significant clay content or for an organic compound) (USEPA, 2001; ISO, 2011). It is recommended that mixtures of spiked sediment be aged for four weeks before starting a test, in keeping with a common practice (USEPA, 1994a, 2000).

[86] Concentrations in sediment are normally calculated and expressed as µg/g or mg/kg, on a dry-weight or wet-weight basis. In some instances, concentrations in pore water might also be measured and expressed as µg/L or mg/L.

Although many studies with spiked sediment have been started within a few hours or days of preparing the mixtures, such short and variable time periods might not be long enough for equilibration of the chemicals mixed in control sediment. A consistent four-week period of aging a mixture before initiating a toxicity test would provide some standardization for intra- and interlaboratory comparisons of results for tests with spiked sediment. Once prepared, each mixture should be placed in a suitable, sealed (with no air space) container, and stored in the dark at 4 ± 2°C (Section 5.2) for four weeks before use.

Based on the objectives of the test, it might be desirable to determine the effect of substrate characteristics (e.g., particle size or organic content) on the toxicity of chemical/sediment mixtures. For instance, the influence of sediment particle size on chemical toxicity could be measured by conducting concurrent multi-concentration tests with a series of mixtures comprised of the test chemical mixed in differing fractions (i.e., segregated particle sizes) or types of natural or formulated control sediment (Section 3.5). Similarly, the degree to which the organic content of sediment can modify chemical toxicity could be examined by performing concurrent multi-concentration tests using different chemical/sediment mixtures prepared with a series of organically enriched control sediments. Each fraction or formulation of natural or artificial control sediment used to prepare such mixtures should be included as a control in the test.

Tests could be required to measure the effect on survival and growth of *H. azteca* at test end, for one or more concentrations of specific chemicals introduced to the test vessel as a solution overlying the sediment. Procedures for preparing test concentrations could vary depending on the objectives of the study. One approach would be to carefully add the test solution(s) to replicate vessels containing a layer of control or other (e.g., field-collected) sediment, with no disturbance or subsequent mixing of the sediment and test solution(s). A second approach would require the test solution(s) introduced to test vessels to be agitated for a predetermined time period in the presence of the sediment before the test organisms are introduced. Chemical/ sediment interactions might differ appreciably depending on the approach taken and could result in a markedly different test result. Unless specified or otherwise required, test water adjusted to 23 ± 1°C should be used to prepare each test solution (Section 6.3). Replicate controls, including solvent controls if a solvent is used, must be prepared and treated identically. Instructions provided earlier in this section on the use of solvents other than water should be followed in preparing solvent controls.

6.3 Test and Control/Dilution Water

The water used for preparing stock or test solutions of chemicals and as test water in 14-day assays with mixtures of spiked sediment should normally be clean test water (see Section 3.4). The source of this water may be reconstituted water or natural water, and might or might not be identical to the water used for culturing the test organisms (see Section 2.3.4). Reconstituted water with a hardness of 120 to 140 mg $CaCO_3$/L (Section 2.3.4; Borgmann, 1996) is recommended if a high degree of standardization is required. For example, the use of a standard reconstituted water is recommended in instances where the measured toxicity of the chemical/sediment mixture is to be compared and assessed relative to toxicity data derived at a number of test facilities for this and/or other chemicals.

6.4 Test Observations and Measurements

A qualitative description of each mixture of spiked sediment and of the overlying test water should be made when the test is being established. This might include observations of the colour, texture, and visual homogeneity of each mixture of spiked sediment, and observations of the colour and opacity of the overlying water. Any change in appearance of the test mixture or overlying water noted during the test, or upon its termination, should be recorded. Measurements of the quality of each mixture of spiked sediment being tested (including the control sediment), and of the overlying water, should be made and recorded as described in Sections 4.5, 5.2, and 5.5.

If analytical capabilities permit, it is recommended that stock solutions, overlying water, sediment, pore water, and test solutions (if studied) be analyzed to determine the chemical concentrations, and to assess whether the sediment has been spiked satisfactorily. In instances where chemical concentrations are to be measured, sample aliquots should be taken from the high, medium, and low test concentrations at the beginning and end of the test, as a minimum. These should be preserved, stored, and analyzed according to suitable, validated procedures.

Unless there is good reason to believe that the chemical measurements are not accurate, toxicity results for any tests in which concentrations are measured should be calculated and expressed in terms of those average measured concentrations determined for both the whole sediment (µg/kg or mg/kg, dry weight) and the pore water (µg/L or mg/L). In cases where concentrations of chemical added to the overlying water are being tested, results should again be expressed as the average measured concentrations determined for the sediment and the pore water, although average chemical concentrations measured for the test solutions overlying sediment should also be calculated and reported (EC, 1992a).

6.5 Test Endpoints and Calculations

Multi-concentration tests with mixtures of spiked sediment are characterized by the 14-day LC50 for the mortality data and the 14-day ICp for the dry weight data. Appropriate statistics and programs for calculating these endpoints are summarized in this section. Section 5.6 provides guidance for calculating and comparing endpoints for single-concentration tests performed with mixtures of spiked sediment. For further information on the appropriate parametric or non-parametric statistics to apply to the endpoint data, the investigator should consult Environment and Climate Change Canada's guidance document on statistical methods for environmental toxicity tests, EPS 1/RM/46 (EC, 2005).

In a multi-concentration test, such as a spiked sediment design, the required statistical endpoints are:

i) an LC50 and its 95% confidence limits for the mortality of *Hyalella*, and

ii) an ICp and its 95% confidence limits for growth (i.e., dry weight of surviving amphipods at test end).

Environment Canada (2005) provides direction and advice for calculating the LC50 and the ICp, including decision flowcharts to guide the selection of appropriate statistical tests. All statistical tests used to derive endpoints require that concentrations be entered as logarithms.

An initial plot of the raw data (percent mortality and dry weight) against the logarithm of concentration is highly recommended, both for a visual representation of the data, and to check for reasonable results by comparison with later statistical computations.[87] Any major disparity between the approximate graphic ICp and the subsequent computer-derived ICp must be resolved. The graph would also show whether a logical relationship was obtained between log concentrations (or, in certain instances, concentration) and effect, a desirable feature of a valid test (EC, 2005).

The optimization of the calculation of the LC50 and its 95% confidence intervals is based on the number of partial effects observed (EC, 2005). In brief, probit and/or logit regression is the preferred method if two partial effects are observed, the *Spearman-Kärber* method is preferred if only one partial effect is observed, and the binomial method is used if no partial effects are observed, and as a general "default" method.

Regression analysis is the principal statistical technique and must be used for the calculation of the ICp,[88,89] provided that the assumptions below are met. A number of models are available to assess dry weight data (using a *quantitative* statistical test) via regression analysis. Use of regression techniques requires that the data meet assumptions of normality and homoscedasticity.

Weighting techniques may be applied to achieve the assumption of homoscedasticity. The data are also assessed for outliers using one of the recommended techniques (see Section 10.2 in EC, 2005). An attempt must be made to fit more than one model to the data. Finally, the model with the best fit[90] must be chosen as the one that is most appropriate for generation of the ICp and associated 95% confidence limits. The Akaike Information Criterion (or an equivalent, such as the Bayesian Information Criterion) is the

[87] As an alternative to plotting the raw data, investigators might choose to calculate and plot the percent inhibition for each test concentration; this calculation is the difference between the average control response and the treatment response (average control response minus average treatment response in the numerator), divided by the average control response (denominator), expressed as a percentage (multiplied by 100%). The value for each treatment is graphed against the concentration; see ASTM (1991a) for more details. The x-axis represents log concentration or, in some instances, concentration, depending on the preferences and purpose of the investigator. For example, using a log scale will match the regression data scales, but concentration might be clearer in the final report. To improve the use of a graph as a visual representation of the data, the investigator might choose to include the regression line as well as the raw data.

[88] The *ICp* is the *inhibiting concentration* for a specified *percent* effect. The "p" represents a fixed *percentage* of reduction, and is chosen by the investigator. Typically, its value is chosen as 25% or 20%.

[89] Historically, investigators have frequently analyzed *quantitative* sublethal endpoints from multi-concentration tests by calculating the *no-observed-effect concentration* (*NOEC*) and the *lowest-observed-effect-concentration* (*LOEC*). Disadvantages of these statistical endpoints include their dependence on the test concentrations chosen and the inability to provide any indication of *precision* (i.e., no 95% or other confidence limits can be derived) (Section 7.1 in EC, 2005). Given these disadvantages, ICp is the required statistical endpoint for dry weight data derived from a multi-concentration test using *Hyalella azteca*.

[90] As described in Section 6.5.8 of EC (2005), Environment Canada's current guidance on statistical methods for environmental toxicity tests specifies the use of the following five models for regression analysis, when estimating the ICp: linear, logistic, Gompertz, exponential and *hormesis* (logistic adapted for hormetic effect at low doses). Specific mathematical expressions of the model, including worked examples for a common statistics package, are also provided in that guidance document (Section 6.5.8 and Appendix O in EC, 2005).

first choice for determining best model fit.[91] Endpoints generated by regression analysis must be bracketed by test concentrations; extrapolation of endpoints beyond the highest test concentration is not an acceptable practice.

If all *Hyalella* in a particular replicate died during the test, there is no measurement of weight, and no entry of data (i.e., it is treated as a missing replicate; as per "Option 1" described in Section 8.2 of EC, 2005). Using this option, mortality would be assessed by a separate analysis. Depending on the objective of the study, however, an investigator might wish to combine the weight endpoint with mortality (i.e., *biomass* or the total weight of all living amphipods divided by the number of juveniles that started in the replicate (see footnote 54 in Section 4.7); as per "Option 3" described in Section 8.3 of EC, 2005).

The ability to mathematically describe *hormesis* (i.e., a stimulatory or "better than the control" response occurring only at low exposure concentrations) in the dose-response curve has been incorporated into recent regression models for quantitative data (see Section 10.3 in EC, 2005). Data exhibiting hormesis can be entered directly, as the model can accommodate and incorporate all data points; there is no trimming of data points that show a hormetic response.

In the event that the data do not lend themselves to regression analysis (i.e., assumptions of normality and homoscedasticity cannot be met), linear interpolation (e.g., ICPIN; see Section 6.4.3 in EC, 2005) can be used to derive an ICp. If the data exhibited hormesis and ICPIN is used, control responses must be entered for those concentrations that demonstrated hormesis (Option 4, Section 10.3.3 in EC, 2005).

For any test that includes a *solvent control sediment* (see Section 6.2), the test results (i.e., survival and final dry weight) for *Hyalella* held in that sediment and in the *clean control sediment* must be examined to determine if they independently meet the test validity criteria (see Section 4.7). If either of these controls fails to meet the test validity criteria, the test results must be considered invalid. If both controls meet the test validity criteria, the results for the two controls must be statistically compared to each other using a *Student's t-test*. If the results for the two controls are not statistically different from each other, then only the data from the *clean control sediment* should be used to calculate the test results.[92] If, however, the survival or final dry weight in the solvent control differs significantly from the results of the clean control sediment, this might be indicative of a potential solvent interference that would then require additional evaluation to determine the impact on the validity of the study. The USEPA (2008) provides guidance on what might be included in such an evaluation:

- assess the relevance of the solvent control response (i.e., percent change relative to the response in control sediment);
- the degree of statistical significance associated with the difference between the two controls (i.e., highly significant difference vs marginally significant difference);

[91] The residual mean square error, previously recommended for this purpose, may be used as a second choice.

[92] The solvent control is not favoured for the calculation of test results by the USEPA because it requires the assumption that the effects of the solvent and toxicant are independent of one another and the current experimental designs do not allow this assumption to be tested (K. Sappington, Office of Pesticide Programs, USEPA, Washington, DC, written communication, 2012).

- assess the breadth of the interference (i.e., are the responses different for both endpoints or just one);

- assess any other potential cause for the interference observed in the solvent control; and

- assess the impact of the potential solvent control interference on uncertainty in the risk estimate.

If a solvent interference is identified, then the solvent control should be used as the basis for calculating results. For each test concentration, including the control treatment(s), the following calculations must be performed and reported:

i) the mean percent survival (± SD) for the *Hyalella*, at the end of the test; and

ii) the mean dry weight ± SD of surviving *Hyalella* at the end of the test.

Section 7

Specific Procedures for 14-Day Water-only Tests

General instructions are given here for testing samples of effluent, leachate, elutriate, receiving water, and chemicals in a 14-day *water-only* survival and growth test using *Hyalella azteca*. These are in addition to the general instructions for testing sediments provided in Section 4. In this section, the word "water" is used for convenience but should be taken to include other similar liquid materials such as effluent, leachate, elutriate, and receiving water.

7.1 General Aspects of the Procedure

The *water-only Hyalella* survival and growth test described in this section is based on two methods:

i) a 10-day water-only method developed by Borgmann *et al*., (2005b) for measuring survival only; and

ii) the 14-day sediment toxicity test (including the 4-day water-only reference toxicity test; Section 4.8), described herein.

The method was developed, and refined (Aquatox Testing and Consulting Inc, 2010; P. Jackman, ALET, Environment and Climate Change Canada, Moncton, NB, personal communication, 2012) to create a tool that could be used alone, or in conjunction with the 14-day *Hyalella* sediment test, to help differentiate between the effects caused by sediment (e.g., historical deposition) and those caused by contaminants in the water column (e.g., receiving water downstream from a wastewater discharge point). These tools are useful for investigation of cause (IOC) studies such as those conducted under the National Environmental Effects Monitoring programme, in which the source of impact is identified. General guidance is given here on the application of the 14-day *Hyalella* survival and growth assay for testing liquid samples in a static-renewal exposure (i.e., renewal three times weekly on non-consecutive days). It can be applied to samples of industrial or sewage effluents, fresh waters (e.g., receiving water), aqueous extracts, or chemical substances that are soluble or that can be maintained as stable suspensions or dispersions under the conditions of the test.

Samples should be tested, following the universal procedures given in Section 4. Additional guidance for testing chemicals is provided in Section 6, and should be followed for testing chemicals in water-only exposures.

7.2 Sample Properties, Labeling, and Storage

General procedures for labelling, transportation, and storing water samples (e.g., effluent, elutriate, leachate, or receiving water) and chemical samples should be as described in Sections 5.2 and 6.1, respectively. For water samples, collapsible polyethylene or polypropylene containers manufactured for transporting drinking water (e.g., Reliance™ plastic containers) are recommended. Their volume can be reduced to fit into a cooler for transport, and the air space within kept to a minimum when portions are removed in the laboratory for the toxicity test or for chemical analyses. Temperature limits for water samples are

those described in Section 5.2, and the samples must not freeze or partially freeze during transport or storage. Upon arrival at the laboratory, the temperature of the sample must be measured and recorded.

Most tests with effluent, leachate, or elutriate will be performed "off-site" in a controlled laboratory facility. Each off-site test must be conducted using one of the following two procedures and approaches:

i) A single sample may be used throughout the test. However, it must be divided into at least three separate containers (i.e., three or more subsamples) upon collection or (in the case of elutriate) preparation. Using this approach, the first subsample must be used for test initiation (Day 0) plus the first two renewals, the second subsample for the 3rd and 4th renewals, and the third subsample for the 5th and 6th renewals.

ii) In instances where the toxicity of the wastewater is known or anticipated to change significantly if stored for up to 10–14 days before use, fresh samples must be collected (or, in the case of elutriate, prepared) on at least three separate occasions using sampling intervals of 4-6 days or less. If three samples are collected at 4- to 6-day intervals (e.g., on Monday and Friday of the first week of testing, and Tuesday or Wednesday of the 2nd week of testing), the first sample must be used for test initiation (Day 0) plus the first two renewals, the second sample for the 3rd and 4th renewals, and the third sample for the 5th and 6th renewals. Wastewaters known or anticipated to be particularly unstable could, if tested off-site, be sampled at daily intervals for 14 consecutive days, and each sample used for only one day of the test in order of sampling.

In those instances where the test is performed on-site in controlled facilities (e.g., within portable or industrial laboratories), fresh samples should be collected and used within 24 h for each replacement of test solution (EC, 2011a).

A sample volume of 60 to 80 L is generally adequate for an off-site multi-concentration test and associated routine sample analyses, using the preceding approach #1. If approach #2 is followed, a per-sample volume (for each of the three samples required to perform the test) of 10 to 15 L should prove adequate in most instances. Greater volumes of effluent would be required if more frequent water renewals (i.e., > 3 times weekly) were conducted. Lesser amounts are required for single-concentration tests (see Section 5).

Testing of effluents, leachates, and receiving waters should commence within 1 day of sampling whenever possible, and must commence no later than 3 days after sampling. Samples of sediment or other solid material collected for extraction and subsequent testing of the elutriate should also be tested as soon as possible (preferably within 1 week), and testing must start no later than 10 days after collection. Procedures given in Environment Canada (1994) for the preparation of elutriates should be followed. Testing of elutriates must commence within 3 days of their preparation, or as specified in a regulation or *protocol*.

7.3 Preparing Test Solutions

Each water sample or subsample in a collection container must be agitated thoroughly just before pouring to ensure the re-suspension of settleable solids and their

homogeneity. Depending on the nature of the sample and the objectives of the test, homogenization of samples might or might not be required before testing. If mixing is carried out, it must be thorough. Subsamples (i.e., a sample divided between two or more containers) must be mixed together to ensure their homogeneity. If further sample storage is required, the composited sample, or a portion of it, should be returned to the subsample containers and stored. The dissolved oxygen content and pH of each sample must be measured just before its use. As necessary, each test solution should be pre-aerated (see Section 7.5.1) before the test solutions are distributed to replicate test vessels.

Filtration of water samples or subsamples is normally not required or recommended. However, if they contain organisms that might be confused with the test organisms, attack them, or compete with them for food, the samples or subsamples must be filtered through a sieve with 60-µm mesh openings, before use (USEPA, 2002). Such filtration could remove suspended solids that are characteristic of the sample or subsample and might otherwise contribute to part of the toxicity or modify the toxicity. In instances where concern exists regarding the effect of this filtration on sample toxicity, a second test should be conducted concurrently using an unfiltered portion of the sample or subsample.

Test solutions of chemicals to be tested are usually prepared by adding aliquots of a stock solution made up in control/dilution water. Alternatively, for strong solutions or large volumes, weighed (using an analytical balance) quantities of chemical may be added to control/dilution water to give the nominal strengths for testing. If stock solutions are used, the concentration and stability of the test chemical in the solution should be determined before the test. Stock solutions subject to photolysis should be shielded from light. Unstable stock solutions must be prepared three times weekly, or as frequently as is necessary to maintain consistent concentrations for each renewal of test solutions. Stock solutions should be prepared as described in Section 6.2. Organic solvents, *emulsifiers*, or *dispersants* should not be used to increase chemical solubility unless it is absolutely necessary or in instances where they might be formulated with the test chemical for its normal commercial purposes. If used, an additional control solution must be prepared containing the same concentration of solubilizing agent that is present in the most concentrated solution of the test chemical. Such agents should be used sparingly and should not exceed 0.1 mL/L in any test solution. If solvents are used, those described in Section 6.2 are preferred.

For any test that is intended to estimate the LC50 as well as the ICp for dry weight (see Section 6.5), at least seven test concentrations plus a control solution (100% *dilution water*) must be prepared, and more (≥8, plus a control) are recommended (as described in Section 6.2 for sediment tests). Each *treatment* including the control(s) must contain a minimum of five replicate test vessels if point-estimates are intended (i.e., LC50 and ICp; see Sections 6.2 and 6.5).

For each definitive test, control solution(s) must be prepared at the same time as the experimental treatments, using an identical number of replicates (see Section 7.4). Any dilution water used to prepare test concentrations must also be used for preparing one set of controls. Each test solution must be mixed well using a glass rod, Teflon™ stir bar, or other device made of non-toxic material. Temperatures must be adjusted as required to 23 ± 1°C.

7.4 Control/Dilution Water

Tests conducted with samples of effluent or leachate or with chemical samples should use, as the control/dilution water, either a supply (source) of the laboratory water shown previously by the testing laboratory to routinely enable valid test results in a 14-day water-only test for survival and growth of *Hyalella* (e.g., natural groundwater, surface water, or reconstituted water), or a sample of the receiving water if there is special interest in a local situation (see Sections 2.3.4 and 3.4). For receiving-water samples collected in the vicinity of a wastewater discharge, chemical spill, or other point-source of possible contamination, "upstream" water may be sampled concurrently and used as control water and diluent for the downstream samples instead of laboratory control water (see Section 3.4 and its associated footnote). The choice of control/dilution water depends on the intent of the test. Because results could be different for the two sources of water, the objectives of the test must be decided before a choice is made (see Sections 5.4 and 6.3). Difficulties and costs associated with the collection and shipment of receiving-water samples for use as control/dilution water should also be considered.

The use of uncontaminated receiving water or an "upstream" water as the control/dilution water can be desirable if site-specific information is required on the potential toxic impact of an effluent, leachate, elutriate, or chemical[93] on a particular receiving water (see Sections 3.4, 5.4, and 6.3). An important example of such a situation would be testing for sublethal effect at the edge of a mixing zone, under site-specific regulatory requirements. Conditions for the collection, transport, and storage of such receiving-water samples should be as described in Section 7.2. Any sample of receiving or "upstream" water used as the control/dilution water for testing effluents, leachates, or receiving waters should be filtered according to the standard recommendation for natural control/dilution water, through a fine mesh net (i.e., ≤60-µm; EC, 2011a). If a sample of receiving or "upstream" water is used as control/dilution water, a separate control solution must be prepared using a laboratory control water (e.g., *test water* or culture water; see Section 3.4). The survival and final dry weight of the *Hyalella* (Section 4.7) in the laboratory control water must be compared to that in the sample of receiving water.

Regardless of the sample type (e.g., water or chemical), tests requiring a high degree of standardization should use reconstituted water as control/dilution water (see Section 3.4). Situations where such use is appropriate include investigative studies intended to interrelate toxicity data for various water types and sources, derived from a number of test facilities or from a single facility where water

[93] Contaminants already in the receiving water might add toxicity to that of the chemical or wastewater being tested. In such cases, uncontaminated dilution water (reconstituted, natural, or dechlorinated municipal) would give a more accurate estimate of the individual toxicity of chemical or material being tested, but not necessarily of the total effect on the site of interest. If the intent of the test is to determine the effect of specific chemical or wastewater on specific receiving water, it does not matter if that receiving water modifies sample toxicity by the presence of additional toxicants, or conversely by the presence of substances that reduce toxic effects, such as humic acids. However, due to the possibility of toxic effects attributable to the "upstream" receiving water, the test must include, as a minimum, a second control using the laboratory's uncontaminated water supply (see Section 3.4). A second series of concentrations using this same water source as the diluents may also be prepared. An alternative (compromise) to using receiving water as dilution and control water is to adjust the pH and hardness of the laboratory water supply (or reconstituted water) to that of the receiving water. Depending on the situation, the adjustment might be to seasonal means, or to values measured in the receiving water at a particular time. Adjustments may be made by the addition or adjustment of the quantities and ratio of reagent grade salts (see Section 2.3.4).

quality is variable. In such instances, it is desirable to minimize any modifying influence due to (differing) dilution-water chemistry.

7.5 Test Conditions

Table 4 provides a checklist of conditions and procedures recommended and required for conducting 14-day water-only toxicity tests using *H. azteca*.

The test procedure, which is largely consistent with the 96-h water-only reference toxicity test described herein (see Section 4.8), uses 2-to 9-day old amphipods, ranging in age by ≤3 days (recommend ≤2 days), to start the test. The test requires 10 organisms per test vessel, and a minimum of 5 replicates per treatment. Recommended test vessels are as described in Section 3.3, with 275 mL solution/vessel. A substrate for the *Hyalella* must be added to each test vessel, and for a given test, the substrate used must be identical for each test solution and each replicate used in that test. Options for test substrate include: an ~3 cm^2 strip of medicinal gauze bandage, presoaked in culture water for 24 h; an ~3 cm^2 piece of Nitex® or plastic mesh (e.g. 500 μm),[94] or a thin layer (i.e., 1–2 mm deep; ~5 mL for the recommended 300 mL high-form glass vessels) of clean silica sand (see footnote 28). The *Hyalella* are transferred to test vessels as described in Section 4.1. Temperature and lighting conditions for this test procedure are the same as those described for sediment toxicity tests and water-only reference toxicity tests (see Sections 4.2 and 4.8; and Tables 2 and 3).

7.5.1 Dissolved Oxygen and Aeration

If (and only if) the measured dissolved oxygen is <40% or >100% of air saturation in one or more test solutions when they have been freshly prepared, each test solution should be pre-aerated before the *Hyalella* are exposed to it. To achieve this, oil-free compressed air should be dispensed through airline tubing and a disposable plastic or glass tube (e.g. capillary tubing or a pipette with an Eppendorf tip, with an opening of about 0.5 mm). The rate of aeration during pre-aeration should not exceed 100 bubbles/min, and the duration of pre-aeration must be the lesser of 20 minutes and attaining 40% saturation in the highest test concentration (or 100% saturation if supersaturation is evident).[95] Any pre-aeration must be discontinued at ≤20 minutes, at which time each test solution should be divided between the replicate test vessels and the test initiated or the solutions used for renewals, regardless of whether the 40 to 100% saturation was achieved in all test solutions.

Solutions in test vessels are normally not aerated during the test, and are covered to minimize evaporation. The required use of oxygen-saturated control/dilution water and the three-times weekly renewal of test solutions are normally adequate in maintaining the dissolved oxygen above the recommended level (i.e., 40%). If the test material or substance has a strong oxygen demand, more frequent renewal of test solutions might be required to maintain DO at ≥40% of saturation. If, however, frequent.

[94] Some chemicals (e.g., PAHs) can be readily adsorbed by nylon mesh, therefore the suspected (or known) toxicants must be taken into consideration when choosing a substrate for water-only tests.

[95] Aeration can strip volatile chemicals from solution or can increase the rate of chemical oxidation and degradation to other substances. However, aeration of test solutions before *Hyalella* exposure might be necessary due to the oxygen demand of the test material (e.g., oxygen depleted in the sample during storage). Aeration also assists in re-mixing the test solution. If it is necessary to aerate any test solution, *all* solutions are to be aerated in the manner stipulated in Section 7.5.1.

Table 4 Checklist of Recommended and Required Conditions and Procedures for Conducting 14-Day Water-only Toxicity Tests Using *Hyalella azteca*

Test type	−	static-renewal 14-day *water-only* toxicity test
Solution replacement	−	three times weekly on non-consecutive days (e.g., Monday, Wednesday, Friday)
Control/dilution water	−	culture water or other clean ground or surface water; *upstream water* to assess toxic impact at a specific location;* reconstituted fresh water if a high degree of standardization is required; natural or reconstituted seawater with salinity ≤15‰ for tests with estuarine sediment; DO, 90 to 100% saturation when used in test
Amphipods	−	removed from known-age culture as <1- to 7-d olds and held in beaker for 2 d preceding test while fed daily; test organisms 2- to 9-d old, and ranging by ≤3 d (recommend ≤2 d) at start of test; 10 animals/test vessel
Substrate for amphipods	−	substrate required; must be identical for all test vessels; options include: a 3 × 3 cm strip of medicinal gauze bandage, presoaked in culture water for 24 h, a 3 × 3 cm piece of Nitex® or plastic mesh, or a 1–2 mm deep (i.e., ~5 ml) layer of clean silica sand
Test vessel	−	glass beaker or glass jar; recommend 300-mL high form ≥7 cm I.D.; normally covered
Volume of test solution	−	275 mL
Number of replicates	−	≥5 test vessels required at each concentration
Number of concentrations	−	minimum of 7, plus control(s); recommend more (i.e., >8), plus control(s)
Temperature	−	daily average, 23 ± 1°C; instantaneous, 23 ± 3°C
Lighting	−	overhead full-spectrum (fluorescent or equivalent); 500 to 1000 lux; 16-h light:8-h dark
Oxygen/Aeration	−	normally none; no pre-aeration unless a test solution has DO <40% or >100% saturation upon preparation, in which case aerate all test solutions for ≤20 minutes at minimal rate before starting test or renewing solution; DO 40 to 100% saturation throughout the test, with more frequent renewal if required to maintain DO; gentle aeration of all vessels only if necessary to meet objectives of test
pH	−	no adjustment if pH of test solutions is in range 6.0 to 8.0;** a second (pH adjusted) test might be required or appropriate, for pH beyond that range

Feeding	–	aqueous suspension of YCT or a 1:1 mixture of YCT and ground commercial fish food flakes (e.g., Nutrafin®, Tetrafin®, or TetraMin®, or Zeigler® Aquatox Feed®), fed daily or three times per week (non-consecutive days); 2.7 mg solids, dry weight added daily to each test vessel if daily feeding; 6.3 mg dry solids added each feeding to each test vessel if fed three times per week only
Observations	–	daily, each vessel, for number of dead or moribund amphipods; mortality and mean dry weight at 14 d
Measurements of	–	temperature daily, representative concentrations; pH, DO, conductivity, salinity (if appropriate) and water quality ammonia at test start and end and before and after each test solution renewal (i.e., minimum of three times weekly) in representative concentrations; recommend hardness and/or alkalinity at test start and end
Endpoints	–	mean percent survival and mean dry weight, each treatment; 14-day LC50 for multi-concentration test, where appropriate; ICp for weight, where appropriate
Reference toxicant	–	copper sulphate ($CuSO_4$), cadmium chloride ($CdCl_2$), potassium chloride (KCl), or sodium chloride (NaCl); 96-hour water-only test for LC50; perform within 14 days of the start of the definitive test; if test organisms are imported, test amphipods from this batch for tolerance to the reference toxicant concurrently with definitive test
Test validity	–	invalid if mean 14-d survival in control water <80%; invalid if average dry weight for replicate control groups at test end is <0.10 mg/amphipod

Effluents, Leachates, and Elutriates

Sample requirement	–	either 3 subsamples from a single sampling or 3 separate samples are collected (or prepared, if elutriate) and handled as indicated in Section 7.2; 60–80 L should be adequate for the assay and routine sample analysis; for on-site tests, fresh samples are collected for each renewal and used within 24 h
Transport and storage	–	if warm (>7°C), must cool to 1 to 7°C with regular ice (not dry ice) or frozen gel packs upon collection; transport in the dark at 1 to 7°C (preferably 4 ± 2°C) using regular ice or frozen gel packs as necessary; sample must not freeze during transit or storage; store in the dark at 4 ± 2°C; use in testing should begin within one day and must start within three days of sample collection or elutriate extraction
Control/dilution water	–	as specified and/or depends on intent; laboratory water or "upstream" receiving water for monitoring and compliance
Suspended solids	–	normally do not filter sample; a parallel test with filtered sample is an option, to assess effects of removed solids

Chemicals

Solvents	–	to be used only in special circumstances; maximum concentration, 0.1 mL/L

Concentration	–	recommend measurements at start and end of each renewal period in aliquots of high, medium, and low strengths and control(s); if concentrations decline 20%, re-test with more frequent renewal (i.e., daily)
Control/dilution water	–	as specified and/or depending on objective; reconstituted water if high degree of standardization required; receiving water if concerned with local toxic impact; otherwise, uncontaminated laboratory water demonstrated to meet test validity criteria
Receiving Water		
Sample requirement	–	as for effluents, leachates, and elutriates
Transport and storage	–	as for effluents, leachates, and elutriates
Control/dilution water	–	as specified and/or depends on intent; laboratory water or if studying local impact use "upstream" receiving water as control/dilution water

* For this option, there must be an additional control using a separate water supply (natural or reconstituted) that has been shown by the testing laboratory to routinely achieve valid test results in previous 14-day water-only tests for survival and growth of *Hyalella azteca*.
** If pH is outside this range, results might reflect toxicity due to biologically adverse pH (see Sections 1.4 and 3.1).

renewal is not successful, and the objectives of the test require DO ≥40% saturation in order to measure toxicity *per se*, then the solution in all test vessels including the controls should be gently aerated (as described in Section 4.3) to maintain appropriate oxygen concentration. Any pre-aeration and/or aeration during the test must be reported including the duration and rate (Section 9).

7.5.2 pH

Toxicity tests should normally be carried out without the adjustments of pH. However, if the sample of test material or substance causes the pH of any test solution to be outside the range of 6.0 to 8.0, results might reflect effects due to pH alone (see Sections 1.4 and 3.1). If it is desired to assess toxic chemical(s) *per se* rather than the deleterious or modifying effects of pH, then the pH of the solutions or sample should be adjusted, or a second, pH-adjusted test should be conducted concurrently using a portion of the sample.[96] For an adjusted test, the initial pH of the sample or of each test solution[97] could, depending on objectives, be adjusted to within ±0.5 pH units of that of the control/dilution water, before exposure of the amphipods. Another acceptable approach for this second, pH-adjusted test is to adjust each test solution, including the control, upwards to pH 6.0 to 6.5 (if the solution has pH <6.0), or downwards to pH 7.5 to 8.0 (if the solution

[96] The usual justification for not adjusting the pH of sample or solution is that pH might have a strong influence on the toxicity of a substance or material being tested. Thus for the (generally) low concentrations of waste found in receiving water after dilution, any change from the natural pH, with concomitant modification of toxicity, should be accepted as part of the pollution "package." That leads to the rationale that the pH should not be adjusted in tests, and that is the requirement to be followed in most instances if test solutions are in the pH range 6.0 to 8.5.

[97] Tests with a chemical, effluent, leachate, elutriate, or aqueous extract of a sediment, which are to receive pH adjustment, might require the separate adjustment of each test solution including the control. Tests with receiving water would normally adjust an aliquot of the undiluted sample, before preparing test concentrations.

has pH >8.0). Solutions of hydrochloric acid (HCl) or sodium hydroxide (NaOH) at strengths ≤1 *N* should normally be used for all pH adjustments. Some situations (e.g., effluent samples with highly buffered pH) might require higher strengths of acid or base.[98]

Abernethy and Westlake (1989) provide useful guidelines for adjusting pH. Aliquots of samples or test solutions receiving pH adjustment should be allowed to equilibrate after each incremental addition of acid or base. The amount of time required for equilibration will depend on the buffering capacity of the solution/sample. For effluent samples, a period of 30 to 60 min is recommended for pH adjustment (Abernethy and Westlake, 1989). For a *Hyalella* test, the adjustment would be made on aliquots used to prepare test concentrations at test initiation and prior to each renewal, the pH in each would be recorded (Section 7.6), and the test started with no further attempt at adjustment.

If the purpose of the toxicity test is to gain an understanding of the nature of the toxicants in the test substance or material, pH adjustment is frequently used as one of a number of techniques (oxidation, filtration, air stripping, addition of chelating agent, etc.) for characterizing and identifying sample toxicity. These "Toxicity Identification Evaluation" (TIE) techniques provide the investigator with useful methods for assessing the physical/chemical nature of the toxicant(s) and its (their) susceptibility to detoxification (USEPA, 1991d, 1992).

7.5.3 Food and Feeding

Organisms in each test vessel must be fed either once daily or three times weekly (on non-consecutive days) throughout the test. An identical food ration must be added to each test vessel on each feeding occasion and the ration provided must be adequate to enable acceptable survival and growth of *H. azteca* during the test period (see Section 7.7), but must not be excessive.[99]

Throughout the test, *H. azteca* are fed either an aqueous mixture of yeast, cereal grass media, and trout chow (YCT; see Appendix H) or a 1:1 mixture of YCT and finely ground commercial fish food flakes (e.g., Nutrafin®, Tetrafin®, TetraMin® or Zeigler® Aquatox Feed; see Section 4.4). If daily feeding is chosen, a ration of 2.7 mg (dry weight) of food or equivalent (i.e., 2.7 mg of ground fish food flakes; or 2.7 mg of YCT; or 1.35 mg of YCT in combination with 1.35 mg fish flakes for the 1:1 mixture of YCT and fish food flakes) must be added daily to each test vessel on Day 0, as well as once per day thereafter until the day the test ends. If the option of feeding 3 times per week is chosen, a ration of 6.3 mg food, dry weight or equivalent (i.e., 6.3 mg ground fish food flakes; or 6.3 mg of YCT; or 3.15 mg of YCT in combination with 3.15 mg fish flakes for the 1:1 mixture of

[98] The rationale for making these adjustments is not really contradictory to the previous rationale of not adjusting pH of wastewaters, but depends on the purpose of the test. Some chemicals and wastewaters will create levels of pH that have direct *sublethal* or *lethal* effects, especially in monitoring or *compliance* tests with full-strength effluent. An investigator might not be primarily interested in whether extreme pH is toxic, because such a pH would be unlikely after even moderate dilution in receiving water, which is naturally well-buffered. If pH *per se* were of primary interest, it could be economically assessed by physicochemical measurements. An investigator would often wish to know if toxic substances were present in a wastewater, and detecting them would require elimination of any masking by toxic action of pH. That rationale leads to the use of pH-adjusted samples or test solutions, in a parallel manner to the standardization of temperature, salinity, and dissolved oxygen at favourable levels when testing for toxic substances.

[99] Excess (i.e., uneaten) food can cause a drop in DO. If this occurs with the daily feeding rate, laboratories should consider the less frequent (i.e., three times weekly) feeding rate. Alternatively, more frequent renewals of the test solutions (i.e., daily) should be considered.

YCT and fish flakes) must be added 3 times per week (starting on Day 0) to each test vessel on non-consecutive days (e.g., on Mondays, Wednesdays, and Fridays) until the day the test ends. Test organisms are not fed on the last day (i.e., Day 14) of the test. Either ration results in the same overall rate of feeding; i.e., 18.9 mg dry food weekly, per test vessel. As described for the sediment test (see Section 4.4), daily feeding might be preferable in order to provide a consistently available fresh food supply; however, excess food can be deleterious (see footnote 99).

7.5.4 Renewal of Test Solutions

This is a static-renewal test where the test solutions are to be almost completely (i.e., ≥80%) renewed a minimum of 3 times weekly, on non-consecutive days throughout the test (i.e., a minimum of 6 times during the 14-day test). Siphoning or use of a pipette is the usual procedure; however, test solutions can be renewed by displacement (i.e., an equivalent volume of water is added 1–2 cm above the bottom of the vessel, and excess water exits through a Nitex® screen at the top of the vessel). If the water is renewed by siphoning, ≤90% of the water is siphoned and replaced for each water renewal. It is desirable to replace solutions in random order across the replicates within a concentration, particularly if the material or substance being tested is difficult to keep mixed because some of the contents settle. During renewal, any uneaten food and other detritus on the bottom of each vessel should be removed. New test solution is slowly added to make up the original total volume of test solution in each vessel. The entire procedure must be done cautiously to prevent any injury or accidental loss of any amphipods. The siphoned or displaced solution should be saved so that an inspection can be made for amphipods that have been accidentally removed. Such amphipods are likely to be injured and should be discarded; the results of the test should be analyzed as if the discarded amphipod had not been present.

7.6 Test Observations and Measurements

Daily observations are made for numbers of dead or moribund amphipods in each test vessel. Temperature must be measured daily in representative test vessels. Extra test vessels may be prepared for the purpose of measuring water temperature during the test. Dissolved oxygen (DO), pH, *conductivity* (and salinity if appropriate) and ammonia must be measured at the start and end of the test, and at each test solution renewal (i.e., minimum of three times weekly). Measurements must be made in both the fresh test solution and the used solution just before or after it is changed in at least the high, medium, and low test concentrations and in the control(s) (see Section 4.5). Hardness and/or alkalinity measurements are recommended for test start and end as described for sediment tests in Section 4.5.

For effluent samples with appreciable solids content, it is desirable to measure total suspended and settleable solids (APHA *et al*., 2005) upon receipt, as part of the overall description of the effluent and as sample characteristics that might influence the results of the toxicity test.

Colour, *turbidity*, odour, and homogeneity (i.e., presence of floatable material or settleable solids) of the sample should be observed at the time of preparing test solutions. *Precipitation*, *flocculation*, colour change, odour, or other reactions upon the preparation of test solutions should be recorded, as should any changes in appearance of solutions during the test (e.g., foaming, settling, flocculation, increase or decrease in turbidity, colour change).

In addition to the observations described above, there are certain additional observations and measurements to be made during tests with chemicals (see Section 6.4). If chemicals are to be measured, sample aliquots should be taken from at least the high, medium, and low test concentrations, and the control(s). As a minimum, separate analyses should be performed with samples taken at the beginning and end of the renewal periods on the first and last days of the test. These should be preserved, stored, and analyzed according to best proven methodologies available for determining the concentration of the particular chemical in aqueous solution.

If chemical measurements indicate that concentrations declined by more than 20% during the test, the toxicity of the chemical should be re-evaluated by a test in which solutions are renewed more frequently than three times weekly (e.g., daily). If necessary, a *flow-through* test could be considered.

All samples should be preserved, stored, and analyzed according to proven methods with acceptable detection limits for determining the concentration of the particular chemical in aqueous solution. Toxicity results for any test in which concentrations are measured should be calculated and expressed in terms of those measured concentrations, unless there is good reason to believe that the chemical measurements are not accurate. In making the calculations, each test solution should be characterized by the geometric average measured concentration to which amphipods were exposed.

The endpoints for the 14-day water-only test are the mean percent survival in each treatment and mean dry weight per surviving amphipod in each treatment measured at the end of the test. Survival and dry weight of amphipods should be assessed at the end of the test as described in Sections 4.7 and 6.5.

7.7 Test Endpoints and Calculations

The test is terminated after 14 days as described in Section 4.6; however, rather than sieving to retrieve the test organisms, all *Hyalella* should be pipetted from the test vessel. The biological endpoints for this 14-day water-only test are survival and dry weight, as described in Section 4.7 for the sediment test. At the end of the 14-day exposure, the following two endpoints must be calculated for each treatment:

i) the mean (± SD) *percentage* of amphipods that survived during the exposure;

ii) the mean (± SD) dry weight per surviving amphipod, calculated from the total weight of the group of survivors.

Missing individuals are assumed to have died and disintegrated during the test, and are included in the tally of dead individuals for a replicate. The total dry weight of the group of survivors in a replicate is then measured.

Toxicity tests might be restricted to a single concentration (e.g. undiluted sample), or might include several concentrations. Statistical analyses and endpoints used for 14-day water-only tests depend on the study objectives and the number of concentrations being tested, and should be consistent with the options and approaches identified in Sections 4.7, 5.6, and 6.5.

Single-concentration tests are often cost-effective for determining the presence or absence of measurable toxicity or as a method for screening a large number of samples for

relative toxicity. Endpoints for these tests would again depend on the objectives of the study, but could include an arbitrary "pass" or "fail" ratings, or percentage mortality at test end. Guidance in Section 5.6 provides instructions that are relevant here, on statistical analysis and reporting of results from a set of tests on different samples, each tested at only one concentration.

Multi-concentration tests are performed in instances where toxicity is anticipated and the test objective is to use a range of concentrations, under standardized conditions, to calculate a point estimate. In these cases the LC50 for survival and the ICp for dry weight must be determined at the end of the test using the guidance for statistical calculations of these endpoints provided in Section 6.5. If a solvent control is included in the study, the guidance on the use of the solvent control data, provided in Section 6.5 must be followed.

The test validity criteria are the same as that described for the 14-day sediment test (Section 4.7). The test is invalid if the average survival for amphipods held in the control water for 14 days is <80% at the end of the test. The test is also invalid if the average dry weight for the replicate control groups is <0.10 mg per individual amphipod surviving at the end of the test.

Reference toxicity tests must be conducted in conjunction with 14-day water-only tests, as described in Section 4.8.

Section 8

Specific Procedures for 42-Day Survival, Growth and Reproduction Test

This section gives specific instructions to be used when conducting a 42-day survival, growth and reproduction test using the freshwater amphipod, *Hyalella azteca*. Instructions for the collection, preparation, and testing samples of control or other sediment spiked experimentally with chemical(s), contaminated sediment, or particulate waste are described. These instructions are in addition to the procedures listed in Sections 4, 5 and 6.

Table 5 provides a checklist of conditions and procedures recommended and required for conducting 42-day survival, growth and reproduction test using *H. azteca*.

The 42-day sediment test is conducted at 23 ± 2 °C as a static-renewal exposure, during which the overlying water is continuously aerated. Overlying water is renewed on Days 14, 28 and 35. The test is conducted in 1 L glass vessels with a sediment to water ratio of 1:50. A minimum of 8 replicates of each sediment sample are required for this test. Each replicate contains 20 amphipods, 7 to 9 days old at the start of the test.[100]

The test organisms are fed (3 times per week on non-consecutive days) a mixture of YCT and ground commercial fish food flakes, and the amount of food provided to the test organisms gradually increases during the 42-d exposure.

In some cases, the 42-day test method will be conducted following evaluation of results from the 14-day survival and growth test. For those samples showing significant growth effects, an evaluation of the more sensitive reproduction endpoint may not be necessary since testing has shown reproduction will not likely occur if average adult weight is <0.35 mg/individual by Day 28 (Ivey *et. al*., 2016).[101]

Test endpoints include mean percentage of adult amphipods that survived during the exposure, mean (± SD) final dry weight of surviving adult amphipods, mean biomass, mean number of young per surviving female per treatment, and mean survival normalized reproduction.

8.1 Sample Collection

General procedures for sample collection are provided in Section 5.1 (field-collected sediment samples) and additional

[100] An additional 2 replicates could be included in order to evaluate test progress (organism growth) at Day 28.

[101] Toxicity tests with *H. azteca* of up to 4 weeks' duration have commonly been performed (e.g., Nebeker et al., 1984; Borgmann and Munawar, 1989; Borgmann and Norwood, 1993; Ingersoll and Nelson, 1990; ASTM, 1991a, 1993; NWRI, 1992). Extending the test duration beyond 14-days enhances the test's ability to discern toxic effects on survival of amphipods; however, the sensitivity of the growth endpoint might not improve due to sexual maturation and associated dimorphism in size of males and females that occurs at this time (Borgmann et al., 1989). Using the conditions and procedures described here, or, as required, appropriate modifications thereof, the test could also be continued beyond 4 weeks to discern and measure effects on reproduction of *H. azteca* (Borgmann et al., 1989; ASTM, 1991a, 1993).

requirements and guidance for the 42-day test method are described herein.

A sediment sample volume of approximately 1 L should be sufficient for conducting the 42-day test, along with sample characterization which includes particle size analysis (percent sand, silt, and clay) and total organic carbon. However, required volumes must be confirmed prior to sample collection, since the exact volume required will depend on the study objectives/design and the nature of the chemical analyses to be performed. For example, the 14-day survival and growth test could be conducted first, followed by evaluation of results and subsequent testing to evaluate sediment effects on reproduction (assuming both tests are started within the 6 week holding time). [102]

As a compromise between logistical and practical constraints (e.g., time and cost) and statistical considerations, only one replicate field sample from each station is required when conducting the 42-d exposure (with each single sample being evaluated using a minimum of 8 laboratory replicates). This approach precludes any determination of mean toxicity at a given sampling location (station), but allows a statistical comparison of toxicity of each sample with the control, and also if desired, a comparison among the test samples (stations), using appropriate statistical tests (see Section 5.6).

One or more sites should be sampled for reference (presumably clean) sediment during each field collection. Sites for collecting *reference sediment* should be sought where the geochemical properties of the sediment are similar to sediment characteristics encountered at the test sites (see Section 5.1).

8.2 Sample Labelling, Transport, Storage, and Analyses

General procedures for sample labelling, transportation, storage and analysis of field-collected sediment samples and spiked sediment samples are provided in Sections 5.2 and 6.1, respectively. Additional requirements and guidance for the 42-day test method are described herein.

Ideally, sediment characteristics that are unstable (e.g., pH, oxidation-reduction potential), changed by conditions of transit and storage (e.g., temperature), or potentially impacted by the exposure conditions (e.g., thin sediment layer) should be measured in the field to help characterize the sample.

Each sample (including all samples of control and reference sediment) must also be characterized by analysing subsamples of whole sediment for particle size distribution (percentage of coarse-grained sand, medium-grained sand, fine-grained sand, silt, and clay) and total organic carbon content.

Other analyses could include (USEPA, 1994a, 2000; APHA *et al.*, 1995, 2010): total inorganic carbon, total volatile solids, biochemical oxygen demand, chemical oxygen demand, cation exchange capacity, acid volatile sulphides, metals, synthetic organic compounds, oil and grease, petroleum hydrocarbons, and pore water analyses for various physicochemical characteristics (e.g., ammonia, pH). Depending on the study objectives, measurement of *redox potential* could be very beneficial when interpreting the exposure route for toxic effects in the 42-day test (i.e., porewater, sediment, water column

[102] Selection of which field replicate to use in the 42-day test will depend to some extent on the study/sampling design, study objectives as well as on the results in each replicate. If no growth effects were observed in any replicate, a randomly selected field replicate could be taken for the 42-day test, or a composite of all the 5 field replicates could be used.

and dietary). The thin sediment layer employed in the 42-d test may not represent accurately the oxic/anoxic conditions normally re-established in the 14-d survival and growth test. Such 'layering' would be present in sediment depths of >3 cm. In comparison, the sediment depth for the 42-d exposure is <1 cm (but still representative of the burrowing depth for *H. azteca*).

8.3 Preparing Samples for Testing

Conditions and procedures for preparing test samples follow the guidance and requirements provided in Section 5.3. For samples evaluated using the 42-day test method, the test volume is sufficiently small that large debris or large indigenous macro-organisms can typically be removed using forceps or a gloved hand before adding sediment to test vessels. If a field-collected sediment contains a large number of indigenous macro-organisms that cannot be removed using forceps, the sample may be press-sieved (not washed) through one or more suitably sized mesh screens. For those sediments containing small macro-organisms that due to sediment characteristics cannot be removed by press-sieving, the sample(s) may be rinsed through a fine-mesh sieve (e.g., 0.25 to 0.5 mm; Day *et al*., 1995b) using any liquid that has separated from the sample during its transport and/or storage. This liquid must be remixed within the sieved sample (Section 4.1). Also see Section 5.3 for additional requirements and guidance.

8.4 Test Water

Overlying water should be from the same source as that used for culturing *H. azteca* (see Sections 2.3.4 and 3.4). Alternatively, this water may be from a separate supply of natural fresh or reconstituted water, or fresh water taken from the reference site near where test sediments were collected. Regardless of source, when used in the 42-day exposure, the test water must contain a minimum bromide (Br) concentration of 0.02 mg/L (added as NaBr) and minimum chloride concentration of 15 mg/L (added as NaCl). In the event Br concentration cannot be confirmed in the overlying water to be used in testing, it must be augmented prior to use to ensure a minimum Br concentration of 0.02 mg/L (also see Section 3.4).

Caution is recommended when considering the use of reconstituted or other overlying water (e.g., site water) for the 42-day survival, growth and reproduction test. Advanced testing would be required to demonstrate that the 42-day test can be successfully conducted (i.e., meet test validity criteria) using reconstituted or other waters during the test. It cannot be assumed that overlying water successfully used for the 14-day survival and growth test would also be acceptable for the longer term 42-day exposure.

If site water is used as overlying water, a second set of controls must be prepared using a supply (source) of laboratory water shown previously by the testing laboratory to routinely enable valid test results in a 42-day test for survival, growth and reproduction of *Hyalella azteca* (see Section 3.4)

Often multiple batches (2-3) of overlying water are prepared for use during the 42-day test (e.g., Batch #1 used at test start, Batch #2 used for Day 14 renewal, and Batch #3 used for Day 28 and 35 renewals). Fewer or more batches can be prepared and used based on number of samples to be evaluated and an individual lab's water quality characteristics. All batches of water must be documented with name of chemical used, amounts used, dates prepared and used.

8.5 Initial Weights

For the 42-day survival, growth and reproduction test, organisms must be ≥7- to ≤9-days old. Starting mean dry weight of the 7 to 9 day old amphipods should be 0.020 – 0.035 mg/individual and should be determined at test initiation. The information on initial weights would provide added confidence that control organisms will achieve adequate growth in the control sediment to reproduce and meet the test validity criteria of ≥6.0 young per surviving female (target should be 8 young per female) by test completion. Method development work has shown reduced reproduction in 42 days when initial mean dry weights are less than 0.020 mg/individual. Measurement of initial weights is strongly recommended for laboratories first using this 42-d exposure.

When conducted, measurement of initial weights will need to be extremely accurate. Because it can be challenging to weigh such small masses, there are several ways of reducing variability between replicates including:

- using the smallest but appropriate size weigh boats;
- ensuring consistency in drying, storage and weighing procedures;
- increasing the number of replicates (beyond the required minimum of 3);
- rinsing organisms with Deionized (DI) or Reverse Osmosis (RO) water to ensure traces of food or other particulates are removed; and
- using a well calibrated balance on a stable surface in a room with minimal disturbance (i.e., not affected by air currents or moisture).

Procedures for obtaining initial dry weights would be similar to those outlined in Section 4.6. In this case, three aluminium weigh boats (previously numbered, dried, weighed, and held in a desiccator) are required for obtaining initial weights of 3 replicates of 10 organisms. Organisms are randomly selected from the batch/group of organisms used for testing (these organisms can be collected before or after organisms are added to test vessels). Due to their small size, care must be taken to ensure the amphipods used for the initial weight measurements are rinsed (e.g., with RO or DI water) until free of sediment or particulates, as even a very small amount of particulate can impact weight measurements. The three weigh boats, each containing 10 organisms are oven dried for 24 h at 60 ± 5 °C. Upon removal from the oven, the boats are moved immediately to a desiccator. Following cooling, each boat should be individually and randomly removed from the desiccator, and weighed immediately to the nearest 10 µg on a balance that measures accurately to this limit. Individual numbers on each weigh boat should also be confirmed to ensure organisms were not lost during the transfers (i.e., from drying oven, to desiccator and balance).

During dry-weight determinations for the groups of amphipods from a test, the first boat weighed should be returned to the desiccator and weighed again at the end as a check on gain of water by the boats in the desiccator to be weighed subsequently. The change should not be >5%; if it is, re-drying of all boats for ≥2 h and reweighing should be carried out. A few weighing boats should be tared, dried, and weighed without amphipods, and results should conform to the laboratory's quality control standards.[103]

[103] Labs may consider developing a control chart of at least 5 trials of 10 weights each using a 20 mg certified weight. This control chart would allow the lab to establish control limits for the balance under various room conditions. On the day initial weights of the test organisms is conducted, 10 weight measurements of the 20 mg certified weight would be conducted and compared to the control chart (T. Watson-Leung, Ontario Ministry of the Environment and Climate Change, personal communication, 2016).

Table 5 Checklist of Recommended and Required Conditions and Procedures for Conducting 42-Day Survival, Growth and Reproduction Test Using *Hyalella azteca*

Universal		
Test type	−	42-day whole sediment toxicity test; static-renewal test
Water renewal	−	Static-renewal of overlying water, at a minimum, on Days 14, 28 and 35
Test (overlying) water	−	culture water or other clean ground or surface water; site water; water adjusted to hardness of site water; reconstituted fresh water for a higher degree of standardization; dissolved oxygen, 90 to 100% saturation when used as overlying water in test; test water must contain a minimum bromide (Br) concentration of 0.02 mg/L (added as NaBr) and minimum chloride concentration of 15 mg/L (added as NaCl); caution recommended when considering the use of reconstituted or other overlying water (e.g., site water); if site water is used as overlying water, a second set of controls must be tested
Acclimation	−	if test water is different from culture water, acclimation of organisms to test water is recommended; acclimation should be conducted on the day preceding the start of the test (Day -1); if imported for use, organisms (7- to 9-days old at test start) must have a minimum acclimation time of 24 hours (at the lab where tests are to be undertaken) prior to use and less than 20% mortality during the 24 hour period immediately preceding the test
Control sediment	−	sample of clean sediment that is used to assess the performance of the test organisms and the acceptability of the test; either natural or formulated sediment
Amphipods	−	test organisms 7- to 9-d old; starting mean dry weight should be 0.020 – 0.035 mg/individual; measurement of initial weights is strongly recommended for laboratories first using the 42-d exposure; 20 amphipods/test vessel
Test vessel	−	wide-mouth 1 L glass beakers or glass jars (e.g., 1 L mason jars approximately 4 inch diameter)
Volume of wet sediment	−	18 mL; sediment to water ratio of 1:50
Volume of test water	−	900 mL; sediment to water ratio of 1:50
Number of replicates	−	must be ≥8 laboratory replicates (10 replicates recommended)
Temperature	−	daily average, 23 ± 2°C; instantaneous, 23 ± 3°C
Lighting	−	overhead full-spectrum (fluorescent or equivalent); 500 to 1000 lux; 16-h light:8-h dark

Aeration	−	continuous and minimal (e.g., 2 to 3 bubbles/s, each test vessel)
Feeding	−	fed (3 times per week on non-consecutive days) a mixture of YCT and ground commercial fish food flakes; amount of food gradually increases during the exposure; based on mg solids (dry weight), food is added to each test vessel as follows: weeks 1 and 2: 3.15 mg YCT and 3.15 mg fish flakes, weeks 3 and 4: 6.3 mg YCT and 6.3 mg fish flakes, and weeks 5 to 6: 12.6 mg YCT and 12.6 mg fish flakes; total feed equals to 264.6 mg solids, dry weight
Observations	−	optional: numbers of amphipods in each vessel seen emerged from sediment, and their behaviour (daily or less frequently); aeration rate and position of pipette to minimize organism and sediment disturbance; amplexus around Day 21
Measurements of overlying water	−	Day 0, in overlying water from at least one test vessel representing each treatment (including control and reference samples): pH, temperature, conductivity, salinity (when appropriate), DO, total ammonia (calculate un-ionized ammonia); also on Day 0, water hardness and alkalinity should be measured in in at least one test vessel representing each treatment; in at least one test vessel representing each treatment: temperature measured daily, conductivity measured weekly, and pH and DO measured three times (3x) per week on non-consecutive days (e.g., Monday, Wednesday and Friday); on Days 7, 21 and 42 in at least one test vessel representing each treatment, measure pH and total ammonia (calculate un-ionized ammonia); on Days 14, 28 and 35 (i.e., renewal days) measure total ammonia in old overlying water, and measure pH and DO in old and fresh solutions; Day 42, in overlying water from at least one test vessel representing each treatment (including control and reference samples): temperature, dissolved oxygen, pH, conductivity, and total ammonia (calculate un-ionized ammonia); also on Day 42, water hardness and alkalinity are recommended measurements in at least one test vessel representing each treatment
Endpoints	−	at test end, document and report number of: surviving adults, male adults, female adults, young; calculate mean (± Standard Deviation, SD) for the following: percentage of adult amphipods that survived during the exposure, final dry weight of surviving adult amphipods, biomass, number of young per surviving female per treatment, and survival normalized reproduction
Test validity	−	invalid if: the average percent survival for adult amphipods held in the control sediment for 42 days is <80% at the end of the test, the average dry weight for the replicate control groups is <0.50 mg per individual adult amphipod surviving at the end of the test, or < 6.0 young are produced per surviving female

Field-collected Sediment or Similar Particulate Material

Transport and storage	−	if sample >7°C, cool to 7°C (ice or frozen gel packs); transport in dark at 1 to 7°C (preferably 4 ± 2°C); store in dark at 4 ± 2°C; test should start within two weeks and must start within six weeks

Reference sediment	–	collected from one or more sampling stations for tests with field-collected sediment; taken from sites presumed to be clean but in the general vicinity of sites where test sediments are collected (i.e., same body of water); frequently selected for use in the toxicity test because of its physicochemical similarity (e.g., particle size and/or organic carbon content) to the test sediments
Sample characterization	–	at least particle size analysis (percent sand, silt, and clay), total organic carbon; other parameters (depending on study objectives) could include: percent water, pore water pH, and pore water ammonia; pore water and sediment sulphide, redox, total inorganic carbon, total volatile solids, biochemical oxygen demand, chemical oxygen demand, cation exchange capacity, total and dissolved metals in sediment and porewater, synthetic organic compounds, oil and grease, petroleum hydrocarbons, Simultaneously Extracted Metals (SEM)/Acid-Volatile Sulfide (AVS) and pore water analyses for various physicochemical characteristics
Preparation of sample	–	only if necessary, remove debris and indigenous macro-organisms using forceps; homogenize sample (including any separated liquid) before the test; if necessary, remove smaller macro-organisms by pressing through fine-mesh sieve (e.g., 0.25 to 0.5 mm), or pass through fine-mesh sieve using liquid that separated from sample during transit and storage and remix this liquid with the sieved sample
Spiked Sediment		
Characterization of chemical(s)	–	information required on stability, water solubility, vapour pressure, purity, and biodegradability should be known for added chemicals spiked into control sediment
Solvent	–	test water is the preferred solvent; if an organic solvent is used, the test must include a solvent control
Preparation of mixtures	–	procedure depends on test design and objectives; might include one or more chemical concentrations mixed in control or test sediment, or specific chemical concentrations added to the test water overlying control sediment; chemical/sediment mixtures may be prepared manually or by mechanical agitation as slurries
Concentration of chemical(s) added	–	normally measure at beginning and end of test, in high, medium, and low strengths as a minimum
Test and dilution water	–	use reconstituted water if a high degree of standardization is required and must meet minimum bromide and chloride concentrations

8.6 Evaluation of Technician Proficiency for Organism Recovery

Technicians experienced with handling, culture and recovery of young amphipods are required for successful conduct of this test.[104] Therefore, before the 42-day toxicity test is undertaken for the first time, an evaluation of technical proficiency must be conducted to ensure staff involved in the tests is trained at recovering and identifying young amphipods. The evaluation involves a 1 day exposure using 2 to 5 day old organisms. The test is conducted in 1 L glass test vessels (e.g., mason jars or beakers) containing 18 mL of wet sediment and 900 mL overlying water, for a sediment to water ratio of 1:50.

Ideally, a variety of sediment types (e.g., high organic matter content (~10%), high sand content, high gravel content) should be used for these trials in order to represent a range of possible test/sediment conditions under which organism recovery would be conducted. A minimum of 3 test vessels per participating technician are required to be evaluated, with a maximum number of 80 amphipods (2-5 days old) per replicate.

The number of technicians participating in the evaluation of technical proficiency must be determined well in advance of the trial in order to ensure sufficient young amphipods are cultured and available for the trial; i.e., a maximum of 240 amphipods (2-5 days old) are required per participating technician.

In this example, 3 technicians evaluate 5 replicate test vessels. On Day 0 the following activities are conducted:

1. 18 mL of control sediment is added to each of 15 x 1 L glass vessels (5 for each participating technician);
2. 900 mL of overlying water is added to each jar and test vessels are labelled with replicate #s 1 through 15;
3. overlying water is gently and continuously aerated overnight, at a temperature of 23 ± 2 °C (aeration and temperature must be continued and maintained until trial termination); and
4. 5 random numbers (e.g., using random number formula in Excel®) are generated with numbers ranging from 20 to 80. This should be done for each participating technician (in this example, 3 sets of 5 random numbers are generated for each of 3 technicians involved in the recovery trial, for a total of 15 random numbers).

On Day 1, the number of organisms distributed to each container will vary (ranging from 20 to 80) and the person dosing the test vessel must be different from those involved in the recovery (so the person recovering the organisms is not aware of the starting number):

1. using the generated random numbers, the required 2-5 day old amphipods are added to each test vessel;
2. ensure technicians assigned to organism recovery are unaware

[104] Results from the inter-laboratory study (Taylor and Novak, 2016) indicated that past experience with *H. azteca* did not guarantee good recovery. This finding supported the need to include an evaluation of technical proficiency in the 42-day test method in order to ensure those involved in the tests are trained at recovering and identifying young amphipods in sediment. Successful recovery (≥ 80%) of young amphipods was possible with minimal experience (e.g., four technicians with experience ranging from 1 to 3 years were able to achieve ≥ 80% recovery). Results also suggested that experience working with young amphipods in a sediment matrix (as opposed to water-only conditions) was important for achieving high organism recovery.

("blind") of the number of organisms added to each container;

3. continue gentle and continuous aeration of overlying water overnight, at a temperature of 23 ± 2 °C; and

4. recovery on Day 1 is an option, with a minimum holding time of 2 hours to allow organisms to burrow into sediment; maximum holding time is overnight, with recovery on Day 2.

After 2-hours, but no later than on Day 2, each technician must record the number of amphipods recovered per replicate. Time to recover organisms is capped at 45-minute per replicate. Typical recovery is 20 to 40 minutes depending on technical experience and sediment characteristics.

Recovery instructions provided in Section 8.13.1 must be followed, and are summarized here:

1. start recovery by removing visible organisms by swirling the test vessel gently and pouring out the overlying water into a clear glass dish (e.g., Pyrex) set on a light table; recovery must be conducted with a light table as a contrasting background, and any dish type that allows the light through to see organisms can be used;

2. all visible amphipods are retrieved with a pipette and counted as removed;

3. using a squirt bottle filled with culture water (or DI/RO water), gently flush the top of the sediment and pour out overlying water/slurry into the clear glass dish; continue recovering and counting organisms in this manner until no sediment remains in the test vessel.

Data from the recovery trials for each technician and replicate are tabulated. Results of organisms recovered are compared to actual number added to ensure mean number of recovered organisms (for each technician conducting the test) is between 80 and 100%. The recommended average recovery (i.e., recovery that labs should strive to achieve) is 85%, and the minimum average recovery must be 80%. A recovery of less than 80% would indicate additional experience and further training is required.[105]

Technical proficiency for organism recovery must be conducted every 3 years (based on calendar year), unless the technician routinely conducts these tests (i.e., 3 times per calendar year).

8.7 Overview of Timing of Events

The 42-day survival, growth and reproduction sediment toxicity test is conducted using a minimum of 8 replicates with each containing 20 amphipods which are 7 to 9 days old at the start of the test. The test is conducted as a static-renewal exposure, with overlying water renewed on Days 14, 28 and 35. An overview of the timing and sequence of test events/activities is provided in Table 6.

[105] Conclusions from a round robin study (Taylor and Novak, 2016) indicated that technicians following the method, using in-house cultures and having some experience working with *H. azteca* can achieve >80% average recovery. For this reason, the recommended average recovery (i.e., recovery that labs should strive to achieve) is 85%, while a minimum of 80% average recovery is required to prove recovery efficiency by technicians prior to performing the test.

Table 6 Overview of Test Activities and Timing of Events for Conducting 42-Day Survival, Growth and Reproduction Toxicity Tests Using *Hyalella Azteca*

Pre-test Activities (completed well in advance of testing)	– conduct "Evaluation of Technical Proficiency for Organism Recovery" – coordinate and schedule culture days to obtain sufficient 7- to 9-day old (at test start) amphipods – if required, prepare overlying water with a minimum bromide concentration of 0.02 mg/L and minimum chloride concentration of 15 mg/L
Day -1 (i.e., 1 day preceding the start of the test)	– if using imported organisms, hold and acclimate in dilution water at test temperature for 24-h – if required, prepare 3 weigh boats for determining the starting mean dry weight of the 7- to 9-day old amphipods – 1abel and rinse test vessels – measure water to be used as overlying water to confirm % DO saturation and pH is within acceptable range for use in testing – homogenize sample(s) and fill a minimum of 8 replicate 1 L glass test vessels (e.g., wide mouth mason jars) with 18 mL of test and control sediment; add 900 mL of control/dilution water as overlying water to each replicate test and control jar – hold test and control vessels with aeration at test conditions (e.g., 23 ± 2 °C, covered, with aeration) overnight
Day 0 (i.e., Beginning the Tests)	– if required, obtain a sub-sample of test organisms to determine starting mean dry weight of the 7- to 9-day old amphipods to be used in testing – monitor water quality (in a representative test vessel per treatment, including control and reference samples), including: o temperature, pH, DO, conductivity, salinity (if appropriate), total ammonia (calculate and report un-ionized ammonia); water hardness and alkalinity should be measured – transfer 20 amphipods (7- to 9-days old; that met culture health validity criteria) to the control and test vessels; organisms are added randomly to the test vessels – release each organism beneath the surface of the water; check test vessels and replace any floating or injured organisms – provide food to each test vessel
Daily During the Test	– measure temperature in a representative test vessel – replenish water loss due to evaporation (if needed using RO or DI water) – gently submerge any floating organisms below the surface of the water using a clean pipette tip – check aeration in all vessels and adjust as required

Weekly During the Test	– measure conductivity in a representative test vessel per treatment, including control and reference samples
Three (3) Times Per Week During the Test On Non-Consecutive Days	– measure pH and DO (before feeding) in a representative test vessel per treatment, including control and reference samples – feed with a gradual increase in amount of food, as follows: o Weeks 1-2: 3.15 mg YCT and 3.15 mg flake fish food o Weeks 3-4: 6.3 mg YCT and 6.3 mg flake fish food o Weeks 5-6: 12.6 mg YCT and 12.6 mg flake fish food o Total feed for 42-day exposure: 264.6 mg solids, dry weight
Day 1	– if required, weigh organisms (3 weigh boats with 10 organisms each) to determine starting mean dry weight of the amphipods used in testing
Day 7	– measure total ammonia and pH in a representative test vessel per treatment, including control and reference samples (calculate un-ionized ammonia) – replenish water loss due to evaporation (use RO or DI water)
Day 14	– measure pH and DO in a representative test vessel per treatment, including control and reference samples (old and fresh solutions) – measure total ammonia (in a representative test vessel per treatment, including control and reference samples) on old solutions (calculate un-ionized ammonia) – renew (~ 80%; 650 to 725 mL) of overlying water
Day 21	– measure total ammonia and pH in a representative test vessel per treatment, including control and reference samples (calculate un-ionized ammonia) – replenish water loss due to evaporation (with RO or DI water)
Day 28	– measure pH and DO in a representative test vessel per treatment, including control and reference samples (old and fresh solutions) – measure total ammonia (in a representative test vessel per treatment, including control and reference samples) on old solutions (calculate un-ionized ammonia) – renew (~ 80%; 650 to 725 mL) of overlying water
Day 35	– measure pH and DO in a representative test vessel per treatment, including control and reference samples (old and fresh solutions) – measure total ammonia (in a representative test vessel per treatment, including control and reference samples) on old solutions (calculate un-ionized ammonia) – renew (~ 80%; 650 to 725 mL) of overlying water
Day 41	– prepare weigh boats for test termination on Day 42

Day 42 (Ending the Exposure)	– do not feed organisms on Day 42 – monitor water quality (in a representative test vessel per treatment, including control and reference samples), including: o temperature, pH, DO, conductivity, total ammonia (calculate and report un-ionized ammonia); water hardness and alkalinity are recommended – at test end, document and report o Number surviving adults o Number male adults o Number female adults o Number young o Place surviving adults on pre-labelled and pre-weighed weigh boats and place in the drying a minimum of 24 hours at 60 ± 5 °C or 6 hours at 100 ± 5 °C.
Day 43	– transfer weigh boats with organisms to desiccator to cool. – once at room temperature, randomly weigh individual weigh boats with organisms
Reported Endpoints	– endpoints reported to include mean % adult survival, mean final dry weight per adult (also as biomass), mean adult growth (weight increase, if initial weights are measured); mean number of young per surviving female
Test Validity Criteria	– ≥80% control adult survival – ≥0.50 mg per adult at 42 days – ≥6.0 young per female

8.8 Beginning the Test

On or before Test Day -1, and if needed prepare three weigh boats for determining the starting mean dry weight of the 7- to 9-day old amphipods (see Section 2.1). Weigh boats can be held in desiccator until ready for use on Day 0 (also see Section 4.6).

Also on Test Day -1, each test vessel placed within the test facility must be clearly coded or labelled to enable identification of the sample or its concentration. The date and time when the test is started must be recorded either directly on the labels or on separate data sheets dedicated to the test (also see Section 4.1).

Following homogenization (see Sections 5.3 and 6.2), collect sample for analysis of required parameters (Section 8.2). Immediately following homogenization, replicate volumes of the sample should be transferred to the 1 L glass test vessels. A minimum of 8 laboratory replicates (10 recommended) with 18 mL of test sediment are required.[106] A minimum of 8 laboratory replicates are also required for each control and reference sediment sample.

[106] A variety of approaches can be used for measurement of 18 mL of sediment; for example, a small container could be calibrated for ~ 18 mL and used to repeatedly fill the test vessels (using a portion of the overlying water to rinse the sediment into the exposure vessels).

Overlying water (900 mL) is added to each replicate test and control vessel to obtain a sediment to water ratio of 1:50. Overlying water is added by slowly pouring the water down the side of the jar. Care must be taken to minimize disturbance of the sediment, but also to ensure any sediment adhering to the sides of test vessel is removed. A mark inscribed on the side of the test vessel can be used to judge the amount of water loss during the test resulting from evaporation (and is used to assess the volume of water to add back).

Test vessels are placed randomly (e.g., out of sequential order) in a 23 ± 2 °C water bath, temperature-controlled room, or temperature controlled cabinet. The space required for the 42-day test will exceed that needed for the 14-day survival and growth test due to the larger size of test vessels (e.g., 1 L mason jars) and therefore, space required for the larger test vessels and longer exposure should be evaluated and confirmed prior to test initiation. Once all test vessels are transferred to the testing area, the top of each jar must be covered. The cover must be transparent (e.g., a plastic petri dish) and can contain a small hole for feeding and another for insertion of a pipette for aeration. Tubing (e.g., Tygon® tubing) is fitted to each pipette tip and connected to aeration gang valves, such that continuous aeration is applied to each test vessel at the rate of 2-3 bubbles per second (also see Section 4.3). In the 42-day test, the pipette tips are suspended ~ 50-60 mm above the sediment (top third of test vessel). Aeration must not disturb the sediment.

Test vessels are to be illuminated with a daily photoperiod of 16-h light and 8-h dark, using overhead, full-spectrum lights (fluorescent or equivalent). Light intensity at the surface of the overlying water should be 500 to 1000 lux. The test system is maintained this way overnight prior to introduction of the organisms.[107]

8.9 Addition of Organisms (Test Day 0)

At the start of the test (Day 0) aeration pipettes (or other appropriate device to deliver air to the test vessels) are temporarily removed and 20 amphipods (7- to 9-day old) are transferred to the test vessels. These organisms should be handled as little and as carefully as possible during their transfer to the test vessels (Section 2.3.9). Amphipods must be placed below the air/water interface in the overlying water. As described in Section 4.1, test organisms may be pipetted directly from a culture vessel into the overlying water, or 20 amphipods may be counted into a transfer vessel (e.g., 30-mL plastic cup) filled with test water at the test temperature, and then recounted before their transfer below the surface of the overlying water. Immediately after test organisms are added, the test vessel is examined and any floating or injured organisms are immediately replaced.

When all organisms have been placed in the test vessels, the start date and time are recorded. Aeration is not resumed until approximately one hour after test initiation to allow organisms to settle in the sediment.

If initial weights are required, 30 organisms (3 replicates of 10 organisms) are also collected for initial dry weight estimates (see

[107] The required minimum equilibration time (overnight) can be extended on a study-specific basis. Equilibration periods of up to 7 days may allow for stabilization of pore water and sediment chemistry (e.g., redox, dissolved metal concentrations) before starting a test. A longer equilibration time may also be a means of reducing labile ammonia. If additional equilibration time is anticipated, additional chemistry data may aid in exposure characterization and data interpretation.

Sections 2.1 and 4.6 for recommendations and procedures for dry weight determination). The organisms used for the initial dry weight measurement must be taken as a sub-sample of the organisms used in testing (i.e., taken randomly from the batch of organisms used, and not from what is remaining after the test is initiated).

8.10 Observations and Measurements During the Test

8.10.1 Chemical Parameters

At the beginning of the test (Day 0) and prior to addition of test organisms, the following must be measured in overlying water from at least one test vessel representing each treatment (including control and reference samples): pH, temperature, conductivity, salinity (if appropriate), DO, and total ammonia. Additional water quality measurements should include alkalinity and water hardness. Un-ionized ammonia must be calculated (see Section 4.5).

As described in Sections 4.5 and 5.5, measurements of DO, conductivity, salinity (if appropriate), pH and temperature can be taken using probes directly inserted into the test vessels. Alternatively, measurements can also be made by extracting an aliquot of the overlying water for analysis. In cases where the sample is not consumed (e.g., DO, pH, conductivity, salinity (if appropriate) and temperature measurements), all water can be returned to the appropriate test vessel and replicate after monitoring is complete. Probes used for water chemistry measurements must be thoroughly rinsed with distilled, deionized or RO water between each measurement to avoid cross contamination of test solutions.

For measurement of water hardness, alkalinity and ammonia requiring permanent removal of overlying water aliquots (i.e., overlying water sample will be consumed for the analysis and not returned to the test vessel), care should be taken to ensure the smallest volume possible is removed from the test replicate (and volume removed must not exceed 10% (90 mL) in any test vessel). Samples of overlying water can also be taken from extra replicates set up for monitoring purposes (or if taken directly from the test vessels, overlying water removed should be replaced with dilution water).

Regardless of method chosen, monitoring of chemical parameters during the test (i.e., after test organisms have been added to the control and test vessels) must be done cautiously to ensure test organisms are not accidentally injured or removed from the test vessels, or adhered to the probes or collection instruments (e.g., pipette tip used to remove water sub-sample). This is particularly important around test day 28, when adults have begun to reproduce and very small, young amphipods can become easily caught in a pipette tip or stuck to a probe.[108]

During the 42-day test, temperature is measured daily and conductivity weekly in at least one test vessel representing each treatment. pH and DO are measured at least three times (3x) per week on non-consecutive days (e.g., Monday, Wednesday and Friday). More frequent (e.g., daily) monitoring of DO should be conducted if declines (e.g., approaching 40% saturation) are observed.

[108] Extra replicates for chemical monitoring purposes are optional for those analyses which require more overlying water (i.e., more than 90 mL) or sediment sub-samples; test organisms might or might not be added to these vessels; in the absence of organism, food might not be added to these additional monitoring test vessels due to fouling concerns in the absence of test organisms (also see Section 5.5).

Total ammonia (with calculated un-ionized ammonia) and pH must be measured on Days 7, 21 and 42 in at least one test vessel representing each treatment.

In addition, on overlying water renewal days (i.e., Days 14, 28 and 35) total ammonia must be measured in the old overlying water just before it is changed. pH and DO must also be measured at the beginning and end of each renewal period, in both the old and fresh overlying water. For old solutions, measurements should be taken from at least one representative test vessel per treatment (control and test vessels). For fresh solutions, measurements should be taken from the composite water sample before adding to each test vessel.

At the end of the test (i.e., Day 42), temperature, dissolved oxygen, pH, conductivity, and total ammonia (calculate un-ionized ammonia) must be measured in at least one test vessel representing each treatment. Measurement of alkalinity and water hardness at test end are also recommneded.

8.11 Overlying Water Replacement and Renewals

Any overlying water lost by evaporation is replaced weekly (minimum) by the addition of temperature-adjusted RO or DI water gently poured down the side of test vessel (a squeeze-bottle can also be used). More frequent replacement can be conducted, if necessary. The water level marked on the testing vessel during the sample preparation (Day -1) is used to evaluate water loss due to evaporation.

On Days 14, 28 and 35, approximately 80% (~ 650 to 725 mL) of the overlying water is renewed. Overlying water renewals may be done manually or automatically (see Sections 3.1 and 4.1). If overlying water is renewed by siphoning, care should be taken to prevent disturbance of the sediment or accidental loss of amphipods emerged from the sediment during this procedure. Particular attention and care is needed once adults have started to produce young (often around Days 24 to 28) such that young are not removed from the test vessel. The end of the siphon must not contact or disturb the sediment.[109]

8.12 Food and Feeding

Critical to the success of the 42-day test, is the provision of an incrementally increasing (doubling) food ration over the course of exposure, in order to accommodate the growth of adults over time and the eventual production of young (Taylor *et al.,* 2016; Ivey *et al.,* 2016). This increased ration and ramping up of the food supply over time allows adults to have two broods that grow to a size that's sufficiently visible within the 42 day test.[110]

Starting on Day 0, organisms in each test vessel must be fed three times weekly (on non-consecutive days (e.g., Monday, Wednesday and Friday)) throughout test. Test organisms are not fed on Day 42. In the event testing is initiated on a Tuesday or Thursday (to meet 3x weekly feeding) a half

[109] Overlying water renewal procedures (used by labs involved in method development) also included gently pouring old/used water into a container and checking to ensure organisms were not lost. Labs using siphoning to remove old/used water also placed a piece of fine mesh over the end of the siphon hose to prevent organism loss. Addition of new/fresh overlying water included gently pouring water down the side of test vessel and use of a diffuser to reduce disruption of the sediment.

[110] Soucek *et al*., 2016 has also shown that control *H. azteca* fed improved diets have their first brood as early as day 24 and the second brood occurring between days 33-35; therefore, the youngest control organisms would be ~ a week old by test completion.

portion of food could be provided on Thursday (Day 0) and half again on Friday (Day 1, the normal feeding) so as to not leave organisms without food. Feeding can then return to full portions the following week (i.e., on Monday, Wednesday and Friday).

The food added to each test vessel is a combination of an aqueous suspension of yeast, cereal grass media and trout chow (YCT) and fish flakes (e.g., TetraMin®). The aqueous mixture of YCT is prepared according to instructions provided Section 4.4 and Appendix H.[111] The fish flake solution is prepared by crushing dry flakes until they pass a #50 sieve (~0.297 mm opening), then adding the ground flakes to DI water and gently shaking. Once added to water, flakes should not be stored longer than 5 days (i.e., it can be prepared on Monday, and used for 3 feedings on Monday, Wednesday and Friday) and must be kept refrigerated.

Based on mg solids (dry weight), food is added to each test vessel as follows:

- Week 1 and 2: 3.15 mg YCT and 3.15 mg fish flakes
- Week 3 and 4: 6.3 mg YCT and 6.3 mg fish flakes
- Week 5 to 6: 12.6 mg YCT and 12.6 mg fish flakes

The total feed equals to 264.6 mg solids, dry weight.

At each feeding, evidence of fungal or plant growth, clarity of the overlying water, or unusual behaviour of the test organisms (e.g., floating) are noted. If water fouling is suspected at any time during the test, DO, pH and ammonia should be measured. This may warrant consideration of increasing the frequency of overlying water renewals beyond the minimum of Days 14, 28 and 35.

[111] YCT can also be purchased, but must ensure that it contains 1.8 g (± 0.1) total solids / L.

8.13 Ending a Test

The test is terminated after 42 days. At least one day prior to test termination, one aluminium weigh boat for each replicate is prepared. Each weigh boat is labelled with the test identification label or number (e.g., control, reference, sample name or number) and replicate number) and then oven-dried for a minimum of 24 hours at 60 ± 5 °C (or 6 hours at 100 ± 5 °C). Weigh boats are transferred to a desiccator to cool and once cooled to room temperature, each weigh boat is individually and randomly weighed on an analytical balance to the nearest 10 µg.

Prior to test termination, two small (e.g., 50 mL) suitable containers (e.g., plastic beaker, petri-dish, weigh boat) for each replicate control and test vessel are also prepared. One container will be used for collection, evaluation, enumeration (and sexing) of adults. The other container will be used for collection, evaluation and enumeration of young amphipods.
Each container is labelled with the test identification label or number (e.g., control, reference, sample name or number) and replicate number.

8.13.1 Recovery of Organisms

On Day 42, adult and young amphipods are recovered from each of the replicate control and test vessels. When starting the recovery process, the contents of each test vessel is gently swirled and one portion of the overlying water is carefully poured out on to a glass dish/pan (e.g., rectangular or round glass Pyrex® dish) placed on a light table.

For successful recovery of young amphipods, a light table must be used as a contrasting background.[112] All visible adult and young amphipods are then removed from the dish. Young are counted as removed using a manual or electronic counter, but should still be collected and placed in the pre-labelled container (in the event recounting is required). Adults are also transferred to their corresponding replicate containers. Individuals which are completely inactive but not obviously dead (e.g., not decomposing) must be examined closely using a low-power microscope or hand-held magnifying glass. These individuals should be prodded gently with a sharp point to confirm that they show no sign of life (i.e., observed for at least 15 seconds), and are then to be counted and recorded as dead and subsequently discarded (i.e., not included in weight measurement).

Even the smallest organisms can usually be detected by this method as there is a very fine dusting of sediment with the water which provides a contrasting background to observe moving organisms. Gentle agitation of the sediment in the recovery dish (either by gently moving the pan/dish or by swirling the sediment using a pipette) can also aide in organism recovery by causing the organisms to move, or by dislodging organisms adhering to sediment particles. In addition, selection of an appropriate dish/pan size, sufficiently large to allow light penetration, is recommended for successful recovery. The use of sieving or staining is not permitted during the recovery process for the 42-day test.

[112] Results from the inter-laboratory study (Taylor and Novak, 2016) supported the need for an adequate contrasting background to be used during the recovery of test organisms. Specifically, when results from those labs not using a light table for recovery were excluded, overall average recovery increased from 76 % (CV = 30%) to 82% (CV = 13%).

Once the first portion of water has been thoroughly examined (and young and adults removed) it can be discarded. Removal of the water/sediment (after being fully examined for amphipods) before additional water/sediment is poured into the recovery dish/pan will help with visibility (e.g., organism will be easier to recover when sediment in the recovery dish is less dark and less dense). Another small amount of the overlying water / sediment slurry is then poured into the glass dish and examined for adult and young amphipods. Additional dilution water can be added to the dish to further dilute the sediment and aide in the recovery process. A squeeze bottle filled with dilution water is used to gently flush the side of jar and the overlying water/slurry is poured into the glass dish. Recovering and counting organisms continues in this manner until no sediment remains in the test vessel.

A consistent amount of time should be taken to examine the contents of each replicate from the control and test vessel. Specifically, 20 to 45 minutes per test vessel should be sufficient to recover >80% of young amphipods produced during the test. Recovery time for each replicate must be documented and held on file. This information could be useful in understanding variability between replicates in hard to process samples.

The total number of young recovered from each replicate is recorded (noting the number of dead young, if any). The total number of live and dead adults recovered from each replicate is also recorded. Unrecovered adults are assumed to be dead and are recorded as such (and are included in statistical analysis).

8.13.2 Sexing of Adults

The number of male and female adults recovered from each replicate must be

identified and recorded.[113] Prior to identification, adults should be free of residual sediment (this can be achieved by a short (~1 minute) rinse with DI or RO water). For each replicate, adult amphipods are then transferred to small containers (e.g., petri-dish, weigh boats) containing a small amount of alcohol (e.g., a few drops of 80% ethanol) or sugar formalin (8% solution of 120 g of sucrose added to 80 mL of formalin and brought to 1 L using de-ionized water; refer to US EPA 2000 for additional details). Each preserved adult amphipod is examined using a dissecting microscope. The sex of each adult is determined by counting the number of adult males with enlarged second gnathopod (see Appendix G).[114]

Other adults (lacking enlarged second gnathopod) are assumed to be females (which are typically smaller in size than the male and may be observed carrying eggs in their brood pouch). A detailed illustration can also be found in US EPA, 2000. Technicans must be trained and able to distinquish male and female amphipods.

After male and female adult amphipods have been identified and enumerated, groups of adults are promptly transferred to their corresponding pre-labelled and pre-weighed aluminium weigh boats. Care must be taken to ensure the amphipods are free of sediment and debris, since even very small amounts of particulates can impact weight measurements. The aluminium weigh boats with organisms are then placed in a drying oven at 60 ± 5 °C for 24 hours. Upon removal from the oven, the boats are moved immediately to a desiccator. Following cooling, each boat should be individually and randomly removed from the desiccator, and weighed immediately to the nearest 10 µg on a balance that measures accurately to this limit. Mean dry weight per adult amphipod that survived the test is calculated for each group.

During the series of dry-weight determinations for the groups of amphipods from a test, the first boat weighed should be replaced in the desiccator and weighed again at the end of all weighings, as a check on gain of water by the boats in the desiccator to be weighed subsequently. The change should not be >5%; if it is, redrying of all boats for ≥2 h and reweighing might be carried out. A few weighing boats should be tared, dried, and weighed without amphipods, and results should conform to the laboratory's quality control standards.

8.14 Test Endpoints and Calculations

The biological endpoints for this 42-day sediment toxicity test are survival, growth (expressed as biomass) and reproduction. Reduced survival, lesser weight and/or reduced reproduction at test end are assessed by comparison with reference and/or control groups (see Sections 5.6 and 6.5). The most sensitive of the three effects is taken as the definitive indication of toxicity.

The following endpoints must be calculated for each treatment in this 42-day sediment toxicity test:

[113] Technicians performing this component of test termination must be able to confidently identify male and female *H. azteca*. Examination of groups of mating pairs from the lab's main culture can be useful in this regard, but should be conducted prior to test termination.

[114] Mating pairs could also be removed first; this might negate the need for ethanol, but pairs must still be examined to confirm sex.

i) mean (± Standard Deviation, SD) percentage of adult amphipods that survived during the exposure[115];

ii) mean (± SD) final dry weight of surviving adult amphipods;

iii) mean (± SD) biomass;

iv) the mean (± SD) number of young per surviving female per treatment; and

v) the mean (± SD) survival normalized reproduction.

The test is invalid if any of the following occurs in the control sediment:

i) the average percent survival for adult amphipods held in the control sediment for 42 days is <80% at the end of the test;

ii) the average dry weight for the replicate control groups is <0.5 mg;

iii) per individual adult amphipod surviving at the end of the test; or

iv) < 6.0 young are produced per surviving female.

The average dry weight per adult amphipod is calculated for the surviving adults in each vessel and must also be calculated as a measurement of biomass. To calculate the biomass endpoint, the total dry weight of the surviving *Hyalella* is divided by the initial number of organisms (20). Currently, this endpoint is used in the fathead minnow larval test in both Canada (EC, 2011a) and the United States (USEPA, 2002), but biomass is not routinely applied to sediment toxicity tests. When making this calculation for biomass, adult *Hyalella* that were accidentally killed, removed or lost during the test, should be deducted from the initial number of organisms for that replicate at the start of the test, as if they had not been in the test. Using final biomass as the statistical endpoint, a value of zero would be assigned if all adults in a particular replicate died during the test (EC, 2005).

The number of young per surviving female is calculated by dividing the total number of young produced in a replicate by the number of adult female organisms surviving at the end of the exposure in that replicate. It is recognized that females dying after reproducing could inflate the apparent reproduction rate when normalized to surviving females; however, correcting for this potential bias is not possible since there is no way to know the sex of organisms that died during the test. For this reason, reproduction data from a replicate that contained zero or one female can be excluded from the calculation of treatment mean. Survival-normalized reproduction is calculated by multiplying the young per surviving female for a replicate by the fraction survival (i.e., males and females) for that replicate.

Statistical analysis of reproduction data must also be considered when conducting the 42-day test. Specifically, a t-test is normally the appropriate method of comparing the data from the test sample to the control sediment or reference sediment. In situations where more than one test sampling station (treatment) is under study, and the investigator wishes to compare multiple sampling stations with the reference, or compare sampling stations with each other, a variety of ANOVA and multiple comparison tests (and non-parametric equivalents) exist.

[115] Missing individuals are assumed to have died and disintegrated during the test, and are included in the tally of dead individuals for a replicate.

Guidance provided in Sections 4.7, 5.6 and 6.5 should be consulted for analysis of the 42-day test results. Additional and more extensive statistical guidance is also available elsewhere (EC, 2005).

Other endpoints that could be measured and/or evaluated include: adult weight increase (requires initial dry weight determination) and bioaccumulation (e.g., Taylor *et al.,* 2016). If initial weights are measured, then the replicate and treatment mean (± SD) weight increase (growth) (including % CV) can be calculated by subtracting the intial average dry weight from the final dry weight of the adults.

Additional details on test design and statistical analysis are provided in Sections 4.7, 5.6 and 6.5) and should be consulted for analysis of the 42-day test results. More extensive statistical guidance is available elsewhere (EC, 2005).

Section 9

Reporting Requirements

Each test-specific report must indicate if there has been any deviation from any of the *must* requirements delineated in Sections 2 to 8 of this biological test method, and, if so, provide details of the deviation. The reader must be able to establish from the test-specific report whether the conditions and procedures preceding and during the test rendered the results valid and acceptable for the use intended.

Section 9.1 provides a list of items that must be included in each test-specific report. A list of items that must either be included in the test-specific report, provided separately in a general report, or held on file for a minimum of five years, is found in Section 9.2. Specific *monitoring* programs, related test *protocols*, or regulations might require selected test-specific items listed in Section 9.2 (e.g., details about the test material and/or explicit procedures and conditions during sample collection, handling, transport, and storage) to be included in the test-specific report, or might relegate certain test-specific information as *data to be held on file*.

Procedures and conditions common to a series of ongoing tests (e.g., routine toxicity tests for monitoring or *compliance* purposes) and consistent with specifications in this document, may be referred to by citation or by attachment of a general report which outlines standard laboratory practice.

Details on the conduct and findings of the test, which are not conveyed by the test-specific report or general report must be kept on file by the laboratory for a minimum of five years so that the appropriate information can be provided if an audit of the test is required. Filed information might include:

- a record of the chain-of-continuity for field-collected or other samples tested for regulatory or monitoring purposes;
- a copy of the record of acquisition for the sample(s) (e.g., courier);
- chemical analytical data on the sample(s) not included in the test-specific report;
- bench sheets for the observations and measurements recorded during the test;
- bench sheets and warning chart(s) for the reference toxicity tests;
- detailed records of the source of the test organisms, their taxonomic confirmation, and all pertinent information on their culturing and health; and
- information on the calibration of equipment and instruments.

Original data sheets must be signed or initialled, and dated by the laboratory personnel conducting the tests.

9.1 Minimum Requirements for a Test-Specific Report

Following is a list of items that must be included in each test-specific report.

9.1.1 Test Substance or Material

- brief description of sample type (e.g., dredged material, reference or contaminated field-collected sediment, control sediment, chemical or chemical substance, effluent, elutriate, leachate, receiving water or liquid extracted from sediments or similar solids), if and as provided to the laboratory personnel;
- information on labelling or coding of each sample;
- date of sample collection; date and time sample(s) received at test facility;
- for wastewater or receiving water samples collected for water-only tests, information on each subsample (i.e., date of collection, date and time subsamples received at the laboratory, and dates or days during test when individual subsamples are used);
- for effluent or leachate for water-only tests, measurement of temperature of sample upon receipt at the test facility;
- for samples or subsamples of wastewater or receiving water collected for water-only tests, measurements of pH and dissolved oxygen, just before its preparation and use in toxicity test; and
- for samples of elutriate or any liquid extracted from sediments or similar solids, dates for sample generation and use; description and procedure for preparation.

9.1.2 Test Organisms

- species and source of brood stock and test organisms;
- range of age, at start of test;
- for the 42-d test, if needed, starting weight range (dry weight) at test initiation (i.e., minimum and maximum weights);
- percentage of young amphipods in *known age* cultures that died or appear to be dead or inactive during the 48-h period immediately preceding the test;
- percentage of young amphipods imported for use in a test that died or appear to be dead or inactive during the 24-h period immediately preceding the test;
- any unusual appearance or treatment of the organisms, before their use in the test; and
- if imported young amphipods are used in the 42-day test, the acclimation period and percent mortality during the 24 hour period immediately preceding the test.

9.1.3 Test Facilities

- name and address of test laboratory;
- name of person(s) performing the test; and
- brief description of test vessels (size and shape).

9.1.4 Test and Control/Dilution Water

- type(s) and source(s) of test and/or control/dilution water;
- measured characteristics of test water, before and/or at time of commencement of toxicity test;

- type and quantity of any chemical(s) added to control or dilution water in water-only tests; and

- for 42-day test, identification if water was augmented with Br and Cl (or analytical confirmation of minimum Br and Cl concentrations in overlying water).

9.1.5 Test Method

- citation of biological test method used (i.e., as per this document);

- for water-only tests, brief description of procedure(s) in those instances in which a sample, subsample, or test solution has been filtered, or adjusted for pH;

- design and description if specialized procedure (e.g., sieving of field-collected test sediment; preparation of mixtures of spiked sediment; preparation and use of solvent and, if so, solvent control; renewal of test solutions at intervals more frequent than three-times weekly; preparation and use of elutriate for water-only tests) or modification of standard test method;

- brief description of frequency and type of observations and all measurements made during test; and

- name and citation of program(s) and methods used for calculating statistical endpoints.

9.1.6 Test Conditions and Procedures

- design and description if any deviation from or exclusion of any of the procedures and conditions specified in this document;

- number of discrete samples per treatment; number of replicate test vessels for each treatment, if applicable; number and description of treatments in each test including the control(s); test concentrations (if applicable);

- depth and volume of sediment and overlying water in each test vessel for sediment tests; volume and depth of test solutions, including controls for water-only tests;

- type of substrate used for water-only tests;

- number of organisms per test vessel and treatment;

- for water-only tests: brief statement (including procedure, rate, and duration) if any pre-aeration of test solutions;

- for spiked sediment tests, time interval between preparation of test sediment and test initiation (i.e., equilibration of test sediment spiked with a chemical as well as equilibration of test sediment with overlying water);

- frequency and rate of overlying water-renewal or indication of static exposure for 14-day sediments tests; frequency and rate of solution renewal for 14-day water-only tests and 42-day sediment tests;

- food type, as well as feeding regime, rate and ration;

- indication of any aeration of overlying water (for sediment tests) or test solutions (for water-only tests),

including rate, prior to and during exposure of test organisms;

- dates when test was started and ended;

- for each sediment sample (including each field replicate and all samples of control and reference sediment): particle size distribution (percentage of coarse-grained sand, medium-grained sand, fine-grained sand, silt, and clay) and total organic carbon content;

- for 14-day sediment tests: all measurements of temperature and dissolved oxygen in overlying water for each treatment made at start of test and three or more times per week thereafter, including test end; all measurements of ammonia and pH for each reference sediment made at start of test and three or more times per week thereafter, including test end; all measurements of conductivity, salinity (if appropriate), pH, and ammonia in overlying water, made at start and end of test for each treatment;

- for water-only tests: all measurements of temperature (daily), as well as pH, DO, conductivity, and ammonia (at test start and before and after each test solution renewal) in test solutions (including controls), made during the test;

- for 42-day test, all measurements in all treatments (control and test) of: temperature (daily), conductivity, salinity (if appropriate) (weekly), dissolved oxygen concentration (as mg/L and % saturation) and pH (three times weekly on non-consecutive days); hardness and alkalinity (if measured) at start and end; total ammonia (and calculated un-ionized ammonia) on days 0, 7, 14, 28, 35, and 42;

- for the 42-day test, on renewal days (Days 14, 28 and 35) measurements of pH, total ammonia and DO at the beginning and end of each renewal period, in both the old and fresh overlying water; and

- date when the reference toxicity test was performed; and description of any deviation from or exclusion(s) of any of the procedures and conditions specified for the reference toxicity test in this document.

9.1.7 Test Results

- for the 14-day sediment and water-only tests: for each replicate (or replicate sample), including each of the control replicates: the number and percentage of mortalities, and the dry weight of surviving amphipods at test end;

- for the 14-day sediment and water-only tests: for each treatment, including controls: mean ± SD for percentage of amphipods that survived the 14-day exposure; mean ± SD for dry weight of surviving amphipods at test end; results of any statistical comparisons;

- for the 42-day tests: number of surviving male and female adults per replicate; mean ± SD percentage of adult amphipods that survived during the exposure (including coefficient of variation (% CV)); replicate and treatment mean ± SD final dry weight (and biomass) of surviving adult amphipods (including % CV); replicate and treatment mean ± SD

weight increase (growth) (including % CV), if initial weights were measured; total number of young per replicate; (including % CV); treatment mean ± SD survival-normalized reproduction; and for each replicate, total number of young per surviving female and mean ± SD number of young per surviving female per treatment (including % CV).

- any LC50 (including the associated 95% confidence limits) and indication of the quantal method used for multi-concentration tests;

- any ICp (together with its 95% confidence limits) determined for the data on dry weight at test end; details regarding any transformation of data that was required, and indication of quantitative statistic used for multi-concentration tests;

- type and results of any statistical analysis performed to determine significant differences between field sampling stations (e.g., logistic regression, contrast analysis, contingency tables);

- type and results of any model fit or significance of parameters tests from logistic regression (if performed);

- any outliers, and justification for their removal or continued inclusion in the data set;

- for a multi-concentration test with spiked sediment, indication as to whether results are based on nominal or measured concentrations of a particular substance or material;

- results for any 96-h LC50 (including its 95% confidence limits) performed with the reference toxicant(s) using the same batch of test organisms, together with the geometric mean value (± 2 SD) for the same reference toxicant(s) as derived at the test facility in previous tests using the procedures and conditions herein; and

- anything unusual about the test, any problems encountered, any remedial measures taken.

9.2 Additional Reporting Requirements

Following is a list of items that must be either included in the test-specific report or the general report, or held on file for a minimum of five years.

9.2.1 Test Substance or Material

- identification of person(s) who collected and/or provided the sample (or subsamples);

- records of sample chain-of-continuity and log-entry sheets; and

- conditions (e.g., temperature, in darkness, in sealed container) of sample (or subsamples) upon receipt and during storage.

9.2.2 Test Organisms

- records of taxonomic confirmation of species, including name of person(s) or facility identifying the organisms and the taxonomic guidelines or method used to confirm species;

- history and age of brood stock;

- description of culture conditions and procedures for *mixed age* and *known*

age cultures, including: facilities and apparatus, lighting, water source and quality, water *pretreatment*, water exchange rate and method, water temperature, age and density in culture, type and quantity of substrate;

- procedures used to count, handle, sort, transfer, and sieve animals; and those to determine their mortality, condition, appearance, and behaviour;

- type, source, and composition of food for cultures and test; records of nutritive value and known contaminants in food; procedures used to prepare and store food, feeding method(s), frequency, and ration; and

- if test organisms are imported (see Section 2.2): records of confirmation of species, by a qualified taxonomist; all supplier's records provided with each shipment, including age and number of test organisms shipped, as well as date and time of shipment; temperature and dissolved oxygen concentration of water in shipment container(s) when shipped and upon arrival.

9.2.3 Test Facilities and Apparatus

- description of systems for regulating lighting and temperature within test facility and for any system providing air and regulating air flow to test vessels;

- description of test vessels, and covers if used;

- description of method and/or apparatus used to deliver and renew overlying water in test vessels; and

- description of procedures used to clean or rinse test apparatus.

9.2.4 Control Sediment, Test Water, and Control/Dilution Water

- for sediment tests: procedures for pretreatment of control sediment (e.g., sieving, settling of sieved fines, formulation and aging if formulated) and test water (e.g., filtration, sterilization, reconstitution and aging if reconstituted, temperature adjustment, aeration rate and duration);

- for water-only tests: details regarding any control/dilution water pre-treatment (e.g., filtration, sterilization, dechlorination; adjustment for pH, temperature, and/or hardness; degassing, aeration rate and duration);

- any ancillary water-quality variables (see Section 2.3.4) measured before and/or during the toxicity test;

- type and quantity of any chemical(s) added to test water or control/dilution water; and

- storage conditions and duration before use, including sampling details if the test water or control/dilution water was "upstream" receiving water.

9.2.5 Test Method

- documentation of "Evaluation of Technical Proficiency for Organism Recovery" demonstrating staff involved in the 42-day test termination are trained and proficient at recovering and identifying young amphipods;

- documentation that technicans conducting the 42-d test are trained

and able to distinquish male and female amphipods;

- description of laboratory's previous experience with this biological test method for measuring sediment toxicity using *H. azteca*;

- procedures used for mixing, spiking, and/or otherwise manipulating test sediments before use;

- procedure used in preparing stock and/or test solutions of chemicals; description and concentration(s) of any solvent used;

- methods used (with citations) for chemical analyses of test material (sediment and pore water), sample, or test solutions including details concerning aliquot sampling, preparation, and storage before analysis; and

- use and description of preliminary or range-finding test.

9.2.6 Test Conditions and Procedures

- photoperiod, light source, and measurements of light intensity adjacent to surface of overlying water or test solution in test vessels;

- records of method of aeration, any disruption of air flow to test vessels during test, and of related DO measurements;

- description of procedure and rate for renewal of overlying water (for sediment tests) or test solutions (for water-only tests);

- appearance of each sample (or mixture thereof), test solution, or of the overlying water (for sediment tests) in test vessels; changes in appearance noted during test;

- for sediment tests: any other chemical measurements (e.g., contaminant concentration, acid volatile sulphides, biochemical oxygen demand, chemical oxygen demand, total inorganic carbon, cation exchange capacity, redox potential, pore water hydrogen sulphide, pore water ammonia) made before and during the test on test material (including control and reference sediment) and contents of test vessels; including analyses of whole sediment, pore water, and overlying water;

- for sediment tests: any other observations or analyses made on the test material (including samples of control or reference sediment); e.g., faunal tracks, qualitative and/or quantitative data regarding indigenous macrofauna or detritus, geochemical analyses;

- for water-only tests: any chemical measurements on sample, stock solutions, or test solutions (e.g., concentrations of one or more specific chemicals; suspended solids content), before and/or at time of the test;

- for 42-day sediment test, confirmation a light table was used for organism recovery;

- conditions, procedures, and frequency for toxicity tests with reference toxicant(s); and

- chemical analyses of concentrations of reference toxicant in test solutions.

9.2.7 Test Results

- results for any range-finding test(s) conducted;

- results for any statistical analyses conducted both with outliers and with outliers removed; for regression analyses, file information indicating sample size (e.g., number of replicates per treatment), parameter estimates with variance or standard error, any ANOVA table(s) generated, plots of fitted and observed values of any models used, results of outlier tests, and results of tests for normality and homoscedasticity;

- warning chart showing the most recent and historic results for toxicity tests with the reference toxicant(s);

- graphical presentation of dose-response data; and

- original bench sheets and other data sheets, signed and dated by the laboratory personnel performing the test and related analyses.

References

Abernethy, S.G. and G.F. Westlake, "Guidelines for pH Adjustment of Effluent Samples for Toxicity Testing," Ontario Ministry of Environment, Rexdale, ON [ISBN No. 0-7729-5974-1], 11p. (1989).

AFNOR (Association française de normalisation), "Détermination de la toxicité des sediments d'eaux douces vis-à-vis de *Hyalella azteca*. Partie 1: Sédiments naturels." Association française de normalisation, Report XP T 90-338-1, September 2003, France (2003).

Ankley, G.T., M.K. Schubauer-Berigan, and J.R. Dierkes, "Predicting the Toxicity of Bulk Sediments to Aquatic Organisms with Aqueous Test Fractions: Pore Water vs. Elutriate," *Environ. Toxicol. Chem.*, 10: 1359-1366 (1991).

Ankley, G.T., D.A. Benoit, R.A. Hoke, E.N. Leonard, C.W. West, G.L. Phipps, V.R. Mattson, and L.A. Anderson, "Development and Evaluation of Test Methods for Benthic Invertebrates and Sediments: Effects of Flow Rate and Feeding on Water Quality and Exposure Conditions," *Arch. Environ. Contam. Toxicol.,* 25: 12-19 (1993a).

Ankley, G.T., V.R. Mattson, E.N. Leonard, C.W. West, and J.L. Bennett, "Predicting the Acute Toxicity of Copper in Freshwater Sediments: Evaluation of the Role of Acid-Volatile Sulfide," *Environ. Toxicol. Chem.,* 12: 315-320 (1993b).

Ankley, G.T., D.A. Benoit, J.C. Balogh, T.B. Reynoldson, K.E. Day, and R.A. Hoke, "Evaluation of Potential Confounding Factors in Sediment Toxicity Tests with Three Freshwater Benthic Invertebrates," *Environ. Toxicol. Chem.,* 13: 1791-1796 (1994).

Ankley, G.T., M.K. Schubauer-Berigan, and P.D. Monson, "Influence of pH and hardness on the toxicity of ammonia to the amphipod *Hyalella azteca*," *Can. J. Fish. Aquat. Sci.* 52: 2078-2083 (1995).

APHA, AWWA, and WEF (American Public Health Association, American Water Works Association, and Water Environment Federation), "Toxicity," Part 8000, in: *Standard Methods for the Examination of Water and Wastewater*, 19th ed., APHA, AWWA, and WEF, Washington, DC (1995).

APHA, AWWA, and WEF (American Public Health Association, American Water Works Association, and Water Environment Federation), "Toxicity," Part 8000, in: *Standard Methods for the Examination of Water and Wastewater*, 21st ed., APHA, AWWA, and WEF, Washington, DC (2005).

AquaTox Testing and Consulting Inc, "Use of a Laboratory-Based Monitoring Tool for Investigation of Cause," Report prepared by Aquatox Testing and Consulting Inc. (Guelph, ON) for Environment Canada, Biological Methods Section. 48p. plus appendices, (2010).

Armstrong, F.A.J. and D.P. Scott, "Photochemical Dechlorination of Water Supply for Fish Tanks with Commercial Water Sterilizers," *J. Fish. Res. Board Can.,* 31: 1881-1885 (1974).

ASTM (American Society for Testing and Materials), "Standard Guide for Conducting Sediment Toxicity Tests with Freshwater Invertebrates," E1383-90, p. 1-20 in: *1991 Annual Book of ASTM Standards – Pesticides, Resource Recovery, Hazardous Substances and Oil Spill Responses, Waste Disposal, Biological Effects, Vol. 11.04*, ASTM, Philadelphia, PA (1991a).

ASTM (American Society for Testing and Materials), "Standard Guide for Collection, Storage, Characterization, and Manipulation of Sediments for Toxicological Testing," E1391-90, p. 1105-1119 in: *1991 Annual Book of ASTM Standards – Pesticides, Resource Recovery, Hazardous Substances and Oil Spill Responses, Waste Disposal, Biological Effects, Vol. 11.04.*, ASTM, Philadelphia, PA (1991b).

ASTM (American Society for Testing and Materials), "Standard Guide for Conducting Sediment Toxicity Tests with Freshwater Invertebrates," E1383-93, p. 1173-1199 in: *1993 Annual Book of ASTM Standards–- Pesticides, Resource Recovery, Hazardous Substances and Oil Spill Responses, Waste Disposal, Biological Effects, Vol. 11.04.*, ASTM, Philadelphia, PA (1993).

ASTM (American Society for Testing and Materials), "Standard Test Methods for Measuring the Toxicity of Sediment-associated Contaminants with Freshwater Invertebrates," E1706-95, p. 1204-1285 in: *1995 Annual Book of ASTM Standards – Pesticides, Resource Recovery, Hazardous Substances and Oil Spill Responses, Waste Disposal, Biological Effects, Vol. 11.05*, ASTM, Philadelphia, PA (1995a).

ASTM (American Society for Testing and Materials), "Standard Guide for Collection, Storage, Characterization, and Manipulation of Sediments for Toxicological Testing," E1706-95, p. 835-855 in: *1995 Annual Book of ASTM Standards– Pesticides, Resource Recovery, Hazardous Substances and Oil Spill Responses, Waste Disposal, Biological Effects, Vol. 11.05,* ASTM, Philadelphia, PA (1995b).

ASTM (American Society for Testing and Materials), "Standard Guide for Use of Lighting in Laboratory Testing," *1999 Annual Book of ASTM Standards*, Vol. 11.05, E-1733-95, p. 1249-1259, American Society for Testing and Materials, Philadelphia, PA (1999).

ASTM. (American Society for Testing and Materials), "Standard Guide for Collection, Storage, Characterization, and Manipulation of Sediments for Toxicological Testing and for the Collection of Samplers Used to Collect Benthic Invertebrates" E1391-03 (2008), in: *2008 Annual Book of ASTM Standards – Biological Effects and Environmental Fate; Biotechnology, Vol. 11.06*, ASTM, Philadelphia, PA 94 p.(2008).

ASTM (American Society for Testing and Materials), "Standard Test Method for Measuring the Toxicity of Sediment-Associated Contaminants with Freshwater Invertebrates," E1706-05 (2010) in: *2010 Annual Book of ASTM Standards – Biological Effects and Environmental Fate; Biotechnology, Vol. 11.06*, ASTM, Philadelphia, PA, 118 p. (2010).

Barnard, J.L. "Marine Fauna of New Zealand: Algal-living Littoral Gammaridea (Crustacea Ampipoda)," *Mem. New Zealand Oceanogr. Inst.,* 62: 7-216 (1972).

Barnard, J.L. and C.M. Barnard, *Freshwater Amphipoda of the World. Vols. I and II,* 830 p., 50 figs., 7 graphs, 98 maps. Hayfield Associates, Mt. Vernon, VA (1983).

Bartlett, A.J., U. Borgmann, D.G. Dixon, S.P. Batchelor, and R.J. Maguire, "Accumulation of Tributyltin in *Hyalella azteca* as an Indicator of Chronic Toxicity: Survival, Growth, and Reproduction," *Environ. Toxicol. Chem.,* 23: 2878-2888 (2004).

Bartlett, A.J., U. Borgmann, D.G. Dixon, S.P. Batchelor and R.J. Maguire, **"**Toxicity and Bioaccumulation of Tributyltin in *Hyalella azteca* from Freshwater Harbour Sediments in the Great Lakes Basin, Canada," *Can. J. Fish. Aquat. Sci.* 62(6): 1243–1253 (2005).

Bartlett, A.J., U. Borgmann, D.G. Dixon, S.P. Batchelor and R.J. Maguire, "Comparison of Toxicity and Bioaccumulation of Tributyltin in *Hyalella azteca* and Five Other Freshwater Invertebrates," *Water Qual. Res. J. Can.,* 42(1): 1-10 (2007).

Bartlett, A.J. and L. Brown, "Is Hyalella Reproduction a Practical Endpoint to Use in Toxicity Testing?" Paper presented at the 38th Annual Aquatic Toxicity Workshop, Winnipeg, MB (2011).

Baudo, R., J.P. Giesy, and H. Muntau, *Sediments: Chemistry and Toxicity of In-place Pollutants*, 405 p., Lewis Publishers, Inc., Chelsea, MI (1990).

Becker, D.S. and G.N. Bigham, "Field Validation of the Amphipod and Chironomid Freshwater Toxicity Tests," SETAC Platform Session #416, *14th Annual Meeting, Soc. Environ. Toxicol. Chem.*, Nov. 14–18, 1993, Houston, TX [Unpublished Report] (1993).

Becker, D.S., G.R. Bilyard, and T.C. Ginn, "Comparisons Between Sediment Bioassays and Alterations of Benthic Macroinvertebrate Assemblages at a Marine Superfund Site: Commencement Bay, Washington," *Environ. Toxicol. Chem.,* 9: 669-685 (1990).

Becker, D.S., C.D. Rose, and G.N. Bigham, "Comparison of the 10-day Freshwater Sediment Toxicity Tests Using *Hyalella azteca* and *Chironomus tentans*," *Environ. Toxicol. Chem.,* 14: 2089-2094 (1995).

Benoit, D.A., G.L. Phipps, and G.T. Ankley, "A Sediment Testing Intermittent Renewal System for the Automated Renewal of Overlying Water in Toxicity Tests with Contaminated Sediments," *Water Res.,* 27(9): 1403-1412 (1993).

Besser, J.M., C.G. Ingersoll, E. Leonard, and D.R. Mount, "Effect of Zeolite on Toxicity of Ammonia in Freshwater Sediments. Implications for Sediment Toxicity Identification Evaluation Procedures." *Environ. Toxicol. Chem.,* 17: 2310-2317 (1998).

Birge, W.J., J. Black, S. Westerman, and P. Francis, "Toxicity of Sediment-Associated Metals to Freshwater Organisms: Biomonitoring Procedures" In: *Fate and Effects of Sediment-Bound Chemicals, Aquatic Systems*, Pergamon Press, NY, pp. 199-218 (1987).

Borgmann, U., "Chronic Toxicity of Ammonia to the Amphipod *Hyalella azteca*; Importance of Ammonium Ion and Water Hardness," *Environ. Pollut.,* 86: 329-335 (1994).

Borgmann, U, "Systematic Analysis of Aqueous Ion Requirements of *Hyalella azteca*: A Standard Artificial Medium Including the Essential Bromide Ion," *Arch. Environ. Contam. Toxicol.* 30: 356-363 (1996).

Borgmann, U., "Toxicity Test Methods and Observations using the Freshwater Amphipod, *Hyalella*," NWRI Contribution No 02-332, National Water Research Institute, Environment Canada, Burlington, ON 50 p., (2002).

Borgmann, U. and A.I. Borgmann, "Control of Ammonia Toxicity to *Hyalella azteca* by Sodium, Potassium and pH," *Environ. Pollut.* 95: 325-331 (1997).

Borgmann, U. and M. Munawar, "A New Standardized Sediment Bioassay Protocol Using the Amphipod *Hyalella azteca* (Saussure)," *Hydrobiologia,* 188/189: 425-531 (1989).

Borgmann, U. and W.P. Norwood, "Spatial and Temporal Variability in Toxicity of Hamilton Harbour Sediments: Evaluation of the *Hyalella azteca* 4-week Chronic Toxicity Test," *J. Great Lakes Res.,* 19(1): 72-82 (1993).

Borgmann, U. and W. Norwood, "Sediment Toxicity Testing Using Large Water-Sediment Ratios: An Alternative to Water Renewal," *Environ. Pollut.* 106: 333-339 (1999).

Borgmann, U. and W. Norwood, "Metal Bioavailability and Toxicity Through a Sediment Core," *Environ. Pollut.* 116: 159-168 (2002).

Borgmann, U., K.M. Ralph, and W.P. Norwood, "Toxicity Test Procedures for *Hyalella azteca*, and Chronic Toxicity of Cadmium and Pentachlorophenol to *H. azteca, Gammarus fasciatus,* and *Daphnia magna*," *Arch. Environ. Contam. Toxicol.,* 18: 756-764 (1989).

Borgmann, U., W.P. Norwood, and K.M. Ralph, "Chronic Toxicity and Bioaccumulation of 2,5,2',5'- and 3,4,3',4'-Tetrachlorobiphenyl and Aroclor 1242 in the Amphipod *Hyalella azteca*," *Arch. Environ. Contam. Toxicol.,* 19: 558-564 (1990).

Borgmann, U., W.P. Norwood, and I.M. Babirad, "Relationship Between Chronic Toxicity and Bioaccumulation of Cadmium in *Hyalella azteca*," *Can. J. Fish. Aquat. Sci.,* 48: 1055-1060 (1991).

Borgmann, U., W.P. Norwood, and C. Clarke, "Accumulation, regulation and toxicity of copper, zinc, lead and mercury in *Hyalella azteca*," *Hydrobiologia,* 259: 79-89 (1993).

Borgmann, U., R. Neron, and W.P. Norwood, "Quantification of Bioavailable Nickel in Sediments and Toxic Thresholds to *Hyalella azteca*," *Environ. Pollut.,* 111: 189-198 (2001a).

Borgmann, U., W. Norwood, T.B. Reynoldson, and F. Rosa, "Identifying Cause in Sediment Assessments: Bioavailability and the Sediment Quality Triad," *Can. J. Fish. Aquat. Sci.,* 58: 950-960 (2001b).

Borgmann, U., M. Nowierski, L.C. Grapentine, and D.G. Dixon, "Assessing the Cause of Impacts on Benthic Organisms Near Rouyn-Noranda, Quebec," *Environ. Pollut.,* 129: 39-48 (2004).

Borgmann, U., W.P. Norwood, and M. Nowierski, "Amphipod (*Hyalella azteca*) solid-phase toxicity test using high water-sediment ratios," In *Small-scale Freshwater Toxicity Investigations*, C. Blaise and J.-F. Férard (eds.), Vol 1, 413-436 (2005a).

Borgmann, U., C.G. Ingersoll, S. Mathyk, and P. Lennie-Misgeld, "Biological Test Method: Test for Survival in 10-Day Water-Only Exposures using the Freshwater Amphipod *Hyalella azteca*, with emphasis on detection of toxicity due to ammonia," Report prepared for the Mackenzie Valley Land and Water Board, Yellowknife, NT (2005b).

Borgmann, U., Y. Couillard, P. Doyle, and D.G. Dixon, "Toxicity of Sixty-Three Metals and Metalloids to *Hyalella azteca* at Two Levels of Water Hardness." *Environ. Toxicol. Chem.,* 24(3): 641-652 (2005c).

Borgmann, U., D.T. Bennie, A.L. Ball, and V. Palabrica, "Effect of a Mixture of Seven Pharmaceuticals on *Hyalella azteca* over Multiple Generations," *Chemosphere,* 66: 1278-1283 (2007).

Borgmann, U., J.E. Schroeder, L.A. Golding, and D.G. Dixon, "Models of Cadmium Accumulation and Toxicity to *Hyalella azteca* During 7- and 28-day Exposures," *Hum. Ecol. Risk Assess.,* 16(3): 560-587 (2010).

Borowsky, B., "The Use of the Males' Gnathopods During Precopulation in Some Gammaridean Amphipoda," *Crustaceana,* 47(3): 245-250 (1984).

Bousfield, E.L., "Freshwater Amphipod Crustaceans of Glaciated North America," *Can. Field-Nat.,* 72(2): 55-113 (1958).

Bousfield, E.L., "Shallow-water Gammaridean Amphipoda of New England," Cornell Univ. Press, Ithaca, NY (1973).

Bousfield, E.L., "A Revised Classification and Phylogeny of the Amphipod Crustacea," *Trans. Roy. Soc. Can.,* 4(14): 343-390 (1979).

Bousfield, E.L., "Amphipoda – Gammaridea," p. 245-285 in: *Synopsis and Classification of Living Organisms*, vol. 2, S.P. Parker, ed., McGraw-Hill, New York, NY (1982).

Bousfield, E.L., "Recent Advances in the Systematics and Biogeography of Landhoppers (Amphipoda: Talitridae) of the Indo-Pacific Region," Proc. Symp. Assoc. Syst. Coll., Honolulu, May 1982, *Bernice P. Bishop Spec. Publ. No. 72,* 171-210 (1984).

Bousfield, E.L., "The Phyletic Classification of Amphipod Crustaceans - Problems in Resolution," Proc. European Crustacean Conference, Paris, September 1992, Oral Presentation and Abstract (1992).

Bousfield, E.L., "A Contribution to the Reclassification of Neotropical Freshwater Hyalellid Amphipods (Crustacea: Gammaridea, Talitroidea)," *Boll. Mus. Civ. St. nat. Verona,* 20: 175-224 (1996).

Bousfield, E.L. and C.P. Staude, "The Impact of J.L. Barnard on North American Pacific Amphipod Research: a Tribute," *Amphipacifica* 1(1): 3-17 (1994).

Brooke, L.T., D.J. Call, G.T. Ankley, D.A. Benoit, C.W. West, and R.A. Hoke, "A Short-Term Method for Estimating the Toxicity of Solid Phase Sediment to the Amphipod *Hyalella azteca*," 51 pp., Report Prepared for the U.S. Environmental Protection Agency, Region 5, Chicago, IL (1993).

Brouwer, H., T. Murphy, and L. McArdle, "A Sediment-contact Bioassay with *Photobacterium phosphoreum*," *Environ. Toxicol. Chem.,* 9: 1353-1358 (1990).

Bulycheva, A.I., "Morskie Blocki Morei SSSR i Sopredel'nych Vod (Amphipoda-Talitroidea)," *Akad. Nauk SSSR, Opred. po Faune SSSR 65* (1957).

Burton, G.A. "Assessing the Toxicity of Freshwater Sediments," *Environ. Toxicol. Chem.,* 10: 1585-1627 (1991).

Burton, G.A., Jr., ed., *Sediment Toxicity Assessment*, 457 p., Lewis Publishers, Inc., Chelsea, MI (1992).

Burton, G.A., Jr. and C.G. Ingersoll, "Evaluating the Toxicity of Sediments," in: *The ARCS Assessment Guidance Document* , Report EPA/905-B94/002, U.S. Environmental Protection Agency, Chicago, IL (1994).

Burton, G.A., Jr., B.L. Stemmer, K.L. Winks, P.E. Ross, and L.A.C. Burnett, "A Multitrophic Level Evaluation of Sediment Toxicity in Waukegan and Indiana Harbors," *Environ. Toxicol. Chem.,* 8: 1057-1066 (1989).

Burton, G.A., Jr., T.J. Norberg-King, C.G. Ingersoll, D.A. Benoit, G.T. Ankley, P.V. Winger, J. Kubitz, J.M. Lazorchak, M.E. Smith, E. Greer, F.J. Dwyer, D.J. Call, K.E. Day, P. Kennedy, and M. Stinson, "Interlaboratory Study of Precision: *Hyalella azteca* and *Chironomus tentans* Freshwater Sediment Toxicity Assays," *Environ. Toxicol. Chem.,* 15: 1335-1343 (1996).

Cairns, M.A., A.V. Nebeker, J.H. Gakstatter, and W.L. Griffis, "Toxicity of Copper-spiked Sediments to Freshwater Invertebrates," *Environ. Toxicol. Chem.,* 3: 435-445 (1984).

Canfield, T.J., N.E. Kemble, W.G. Brumbaugh, F.J. Dwyer, C.G. Ingersoll, and J.F. Fairchild, "Use of Benthic Invertebrate Community Structure and the Sediment Quality Triad to Evaluate Metal-contaminated Sediment in the Upper Clark Fork River, MT," *Environ. Toxicol. Chem.,* 13: 1999-2012 (1994).

Cano, M.L., S.D. Dyer, and A.J. DeCarvalho, "Effect of Sediment Organic Carbon on the Toxicity of a Surfactant to *Hyalella azteca*," *Environ. Toxicol. Chem.,* 15: 1411-1417 (1996).

CCME (Canadian Council of Ministers of the Environment), "Canadian Water Quality Guidelines for the Protection of Aquatic Life: Reactive Chlorine Species" Canadian Council of Ministers of the Environment, Winnipeg, 9 p (1999).

Chapman, P.M., R.N. Dexter, S.F. Cross, and D.G. Mitchell, "A Field Trial of the Sediment Quality Triad in San Francisco Bay," NOAA Tech. Mem. NOS OMA 25, National Oceanic and Atmospheric Administration, Rockville, MD (1986).

Chapman, P.M., R.N. Dexter, and E.R. Long, "Synoptic Measures of Sediment Contamination, Toxicity and Infaunal Community Composition (the Sediment Quality Triad) in San Francisco Bay," *Marine Ecol.,* 37: 75-96 (1987).

Chapman, P.M., E.A. Power, R.N. Dexter, and H.B. Andersen, "Evaluation of Effects Associated with an Oil Platform, Using the Sediment Quality Triad," *Environ. Toxicol. Chem.,* 10: 407-424 (1991).

Conlan, K.E., "Revision of the Crustacean Amphipod Genus *Jassa* Leach (Corophioidea: Ischyroceridae)," *Can. J. Zool.,* 86: 2031-2075 (1990).

Conlan, K.E., "Precopulatory Mating Behaviour and Sexual Dimorphism in the Amphipod Crustacea," *Hydrobiologica,* 223: 255-282 (1991).

Cooper, W.E. "Dynamics and Productivity of a Natural Population of a Freshwater Amphipod," *Hyalella azteca*, *Ecol. Monogr.,* 35: 377-394 (1965).

Day, K.E., B.J. Dutka, K.K. Kwan, N. Batista, T.B. Reynoldson, and J.L. Metcalfe-Smith, "Correlations Between Solid-phase Microbial Screening Assays, Whole-sediment Toxicity Tests with Macroinvertebrates and *In-situ* Benthic Community Structure," *J. Great Lakes Res.,* 21: 192-206 (1995a).

Day, K.E., S. Kirby, and T.B. Reynoldson, "The Effect of Manipulations of Freshwater Sediments on Responses of Benthic Invertebrates in Whole-sediment Toxicity Tests," *Environ. Toxicol. Chem.,* 14: 1333-1343 (1995b).

de Groot, A.J. and K.H. Zschuppe, "Contribution to the Standardization of the Methods of Analysis for Heavy Metals in Sediments," *Rapp. P.-v. Reun. cons. Int. Explor. Mer.,* 181: 111-122 (1981).

Deitzer G. "Spectral Comparisons of Sunlight and Different Lamps," p. 197-199 in: Proceedings of International Lighting in Controlled Environments Workshop (Tibbits, T.W., ed), 1 Mar. 1994, Madison, WI (1994).

de March, B.G.E., "The Effects of Photoperiod and Temperature on the Induction and Termination of Reproductive Resting Stage in the Freshwater Amphipod *Hyalella azteca* (Saussure)," *Canad. J. Zool.,* 55: 1595-1600 (1977).

de March, B.G.E., "The Effects of Constant and Variable Temperatures on the Size, Growth and Reproduction of the Freshwater Amphipod *Hyalella azteca* (Saussure)," *Can. J. Zool.,* 56(8): 1801-1806 (1978).

de March, B.G.E., "Survival of *Hyalella azteca* (Saussure) Raised Under Different Laboratory Conditions in a pH Bioassay, with References to Copper Toxicity," *Can. Fish. Mar. Serv. Tech. Rep. No. 892* (1979).

de March, B.G.E., "*Hyalella azteca* (Saussure)," p. 61-77 in: *Manual for the Culture of Selected Freshwater Invertebrates*, Canad. Spec. Publ. Fish. Aquat. Sci. No. 54, Department of Fisheries and Oceans, Ottawa, ON (1981).

Denton, D., J. Diamond, and L. Zheng, "Test of Significant Toxicity: A Statistical Application for Assessing Whether an Effluent or Site Water is Truly Toxic," *Environ. Toxicol. Chem.,* 30: 1117-1126 (2011).

DFO (Department of Fisheries and Oceans), "Bioassay Procedures for *Hyalella azteca*," Standard Operating Procedure, U. Borgmann and W.P. Norwood, 3 p., May 11, 1992, Great Lakes Laboratory for Fisheries and Aquatic Sciences, Burlington, ON [Unpublished Report] (1992).[116]

Di Toro, D.M., J.H. Mahony, D.J. Hansen, K.J. Scott, M.B. Hicks, S.M. Mayr, and M. Redmond, "Toxicity of Cadmium in Sediments: the Role of Acid Volatile Sulfides," *Environ. Toxicol. Chem.,* 9: 1487-1502 (1990).

Di Toro, D.M., C.S. Zarba, D.J. Hansen, W.J. Berry, R.C. Swartz, C.E. Cowan, S.P. Pavlou, H.E. Allen, N.A. Thomas, and P.R. Paquin, "Technical Basis for Establishing Sediment Quality Criteria for Nonionic Organic Chemicals Using Equilibrium Partitioning," *Environ. Toxicol. Chem.,* 10: 1541-1583 (1991).

Ditsworth, G.R., D.W. Schults, and J.K.P. Jones, "Preparation of Benthic Substrates for Sediment Toxicity Testing," *Environ. Toxicol. Chem.,* 9: 1523-1529 (1990).

Doe, K.G. and S. Wade, "Effects of Temperature and Salinity on *Amphiporeia virginiana* During Ten-day Sediment Bioassays," Unpublished Report Prepared by Atlantic Region Laboratory Division, Environment Canada, Dartmouth, NS (1992).

[116] Procedures are the same as Borgmann *et al.* (1989) for culturing, Borgmann *et al.* (1991) for water-only tests, and Borgmann and Norwood (1993) for sediment tests.

Dunnett, C.W., "A Multiple Comparison Procedure for Comparing Several Treatments with a Control," *J. Amer. Stat. Assoc.,* 50: 1096-1121 (1955).

Dwyer, F.J., C.J. Ingersoll, and N.E. Kemble, "Use of Standardized Formulated Sediment in Toxicity Tests," SETAC Platform Session #275, *14th Annual Meeting, Soc. Environ. Toxicol. Chem*., Nov. 14–18, 1993, Houston, TX [Unpublished Report] (1993).

EC (Environment Canada), "Guidance Document on the Control of Toxicity Test Precision Using Reference Toxicants," Conservation and Protection, Ottawa, ON, Report EPS 1/RM/12, 85 p. (1990).

EC (Environment Canada),"Biological Test Method: Acute Test for Sediment Toxicity Using Marine or Estuarine Amphipods," Conservation and Protection., Ottawa, ON, Report EPS 1/RM/26, 83 p. (1992a).

EC (Environment Canada), "Biological Test Method: Test of Reproduction and Survival Using the Cladoceran *Ceriodaphnia dubia*," Conservation and Protection, Ottawa, ON, Report EPS 1/RM/21, 73 p. (1992b).

EC (Environment Canada), "Biological Test Method: Test of Larval Growth and Survival Using Fathead Minnows," Conservation and Protection, Ottawa, ON, Report EPS 1/RM/22, 70 p. (1992c).

EC (Environment Canada), "Guidance for the Collection and Preparation of Sediments for Physicochemical Characterization and Biological Testing," Environmental Protection Service, Ottawa, ON, Report EPS 1/RM/29, 144 p. (1994).

EC (Environment Canada), "Guidance Document on Measurement of Toxicity Test Precision Using Control Sediments Spiked with a Reference Toxicant," Environmental Protection Service, Ottawa, ON, Report EPS 1/RM/30, 56 p. (1995).

EC (Environment Canada), "Test for Survival and Growth in Sediment Using Larvae of Freshwater Midges (*Chironomus tentans* or *Chironomus riparius*)," Environmental Protection Service, Ottawa, ON, Report EPS 1/RM/32 (1997a).

EC (Environment Canada), "Test for Survival and Growth in Sediment Using the Freshwater Amphipod *Hyalella azteca*," Environmental Protection Service, Ottawa, ON, Report EPS 1/RM/33 (1997b).

EC (Environment Canada), "Recommended Procedure for Adjusting Salinity of Effluent Samples for Marine Sublethal Toxicity Testing Conducted Under Environmental Effects Monitoring (EEM Programs)," Unpublished Report, July 1997, 5 p., Method Development and Application Section, Environmental Technol. Centre, Ottawa, ON (1997c).

EC (Environment Canada), "Guidance Document on the Application and Interpretation of Single-Species Tests in Environmental Toxicology," Environmental Protection Service, Ottawa, ON, Report EPS 1/RM/34 (1999a).

EC (Environment Canada), "Recommended Procedure for the Importation of Test Organisms for Sublethal Toxicity Testing," Method Development and Applications Section, Environmental Technology Centre, Ottawa, ON, unpublished report, Sept. 1999, 22p. (1999b).

EC (Environment Canada), "Guidance Document on Statistical Methods for Environmental Toxicity Tests" including 2007 amendments, Environmental Protection Service, Ottawa, ON, Report EPS 1/RM/46, 241 p. (2005).

EC (Environment Canada), "Procedure for pH Stabilizaion During the Testing of Acute Lethality of Wastewater Effluent to Rainbow Trout," Environmental Protection Service, Ottawa, ON, Report EPS 1/RM/50, 21 p. (2008).

EC (Environment Canada), "2010 Pulp and Paper Environmental Effects Monitoring (EEM) Technical Guidance Document," Environment Canada, 490 p. (2010).

EC (Environment Canada), "Biological Test Method: Test of Larval Growth and Survival Using Fathead Minnows," Environment Canada, Environmental Protection Service, Ottawa, Ontario, Report EPS 1/RM/22, second edition, 73 p. (2011a).

EC (Environment Canada), "Metal Mining Environmental Effects Monitoring (EEM) Technical Guidance Document," Environment Canada http://www.ec.gc.ca/esee-eem/default.asp?lang=En&n=4866BC23-1, In Preparation (2011b).

EC (Environment Canada), "Test for Survival and Growth in Sediment and Water Using the Freshwater Amphipod *Hyalella azteca*," Environmental Protection Service, Ottawa, ON, Report EPS 1/RM/33, Second Edition (2013).

FDA (Food and Drug Administration), "*Hyalella azteca* Acute Toxicity," Environmental Assessment Technical Assistance Document 4.10, 22 p., March 1987, Prepared by C. Eirkson, M.C. Harrass, C.M. Osborne, P.G. Sayre, and M. Zeeman., Center for Food Safety and Applied Nutrition and Center for Veterinary Medicine, Washington, DC (1987).

Hamr, P., R.S. Kirby, P. Gillis, and K.E. Day, "Development of Methodologies for Whole-sediment Toxicity Tests with Benthic Invertebrates," Unpublished Report TS-27, Prepared for Environment Canada by National Water Research Institute, Burlington, ON (1994).

Hargrave, B.T., "The Utilization of Benthic Microflora by *Hyalella azteca*," *J. Animal Ecology,* 39: 427-437 (1970).

Hockett, J.R., T.L. Highland, D.J. Hoff, D.R. Mount, T.J. Norberg-King, and T. Valenti, "Modifying Foods and Feeding Regimes to Optimize the Performance of *Hualella azteca* During Chronic Toxicity Tests," SETAC – 32nd North American Annual Meeting, Nobember 2011, Boston, MA (2011).

Hoenig, J.M., and D.M. Heisey, "The Abuse of Power: The Pervasive Fallacy of Power Calculations for Data Analysis," *Amer. Statistician,* 55: 19–24 (2001).

Hoke, R.A., P.A. Kosian, G.T. Ankley, A.M. Cotter, F.M. Vandermeiden, G.L. Phipps, and E.J. Durhan, "Check Studies with *Hyalella azteca* and *Chironomus tentans* in Support of the Development of a Sediment Quality Criterion for Dieldrin," *Environ. Toxicol. Chem.,* 14: 435-443 (1995).

Ingersoll, C., "Desirable and Necessary Attributes for Freshwater Sediment Toxicity Tests: *Hyalella azteca*," p. 269-280 in: *Proceedings, Tiered Testing Issues for Freshwater and Marine Sediments*, September 16-18, 1992, USEPA, Washington, DC (1992).

Ingersoll, C.G. and M.K. Nelson, "Testing Sediment Toxicity with *Hyalella azteca* (Amphipoda) and *Chironomus riparius* (Diptera)," p. 93-109 in: *Aquatic Toxicology and Risk Assessment*, W.G. Landis and W.H. van der Schalle (eds). 13th Volume, ASTM STP 1096, ASTM, Philadelphia, PA (1990).

Ingersoll, C.G., F.J. Dwyer, S.A. Burch, M.K. Nelson, D.R. Buckler, and J.B. Hunn, "The Use of Freshwater and Saltwater Animals to Distinguish Between the Toxic Effects of Salinity and Contaminants in Irrigation Drain Water," *Environ. Toxicol. Chem.,* 11: 503-511 (1992).

Ingersoll, C.G., G.T. Ankley, D.A. Benoit, E.L. Brunson, G.A. Burton, F.J. Dwyer, R.A. Hoke, P.F. Landrum, T.J. Norberg-King, and P.V. Winger, "Toxicity and Bioaccumulation of Sediment-associated Contaminants Using Freshwater Invertebrates: A Review of Methods and Applications," *Environ. Toxicol. Chem.,* 14: 1885-1894 (1995).

ISO (International Organization for Standardization), "Water quality -- Determination of the Toxic Effect of Sediment and Soil Samples on Growth, Fertility and Reproduction of *Caenorhabditis elegans* (Nematoda)," ISO 10872:2010, Geneva, Switzerland (2010).

ISO (International Organization for Standardization), "Water Quality-Determination of Toxicity of Fresh Water Sediments Using *Hyalella azteca*" Report No. ISO 16303: 2013 Geneva, Switzerland.

Ivey, C, C. Ingersoll, and E. Brunson, "An Evaluation of Reconstituted Water for Use in Chronic Water-Only and Sediment Toxicity Tests with *Hyalella azteca*," SETAC – 25th North American Annual Meeting, November 2004, Portland OR (2004).

Ivey, C., C. Ingersoll, N. Kemble, J. Kunz, D.R. Mount, and J.R. Hockett, "Evaluation of the Influence of Bromide or Iodide on the Performance the Amphipod *Hyalella azteca* in Reconstituted Waters," SETAC – 32nd North American Annual Meeting, Nobember 2011, Boston, MA (2011).

Ivey C.D., C.G. Ingersoll, B.G. Brumbaugh, E. Hammer, D.R. Mount, R. Hockett R, T. Norberg-King, D.J. Soucek and L.N. Taylor, "Using an Inter-laboratory Study to Revise Methods for Conducting 10- to 42-d Water or Sediment Toxicity Tests with *Hyalella azteca,*" *Environ. Toxicol. Chem.*, 35: 2439-2447 (2016).

Jackman, P.M. and K.G. Doe, "Laboratory Studies of Confounding Factors Affecting Sediment Toxicity Tests and Integration of These Results in Interpretation of Field Studies," Paper presented at the 27th Annual Aquatic Toxicity Workshop, St. John's, NF (2000).

Kemble, N.E., W.G. Brumbaugh, E.L. Brunson, F.J. Dwyer, C.G. Ingersoll, D.P. Monda, and D.F. Woodward, "Toxicity of Metal-contaminated Sediments from the Upper Clark Fork River, MT to Aquatic Invertebrates in Laboratory Exposures," *Environ. Toxicol. Chem.,* 13: 1985-1997 (1994).

Kubitz, J.A. "Protocol for Small Volume Bioassay for *Hyalella azteca* - Survival Endpoint," Method Standard Operating Procedure Check List, Draft Release 1.0, 6 p., February 9, 1993. Institute for Environmental Toxicology, Michigan State University, East Lansing, MI [Unpublished Report] (1993a).

Kubitz, J.A. "Culture of the Amphipod *Hyalella azteca* (Saussure)," Method Standard Operating Procedure Check List, Draft Release 1.1, 8 p., April 14, 1993. Institute for Environmental Toxicology, Michigan State University, East Lansing, MI [Unpublished Report] (1993b).

Kubitz, J.A., J.M. Besser, and J.P. Giesy, "Optimizing the Experimental Design of a Sediment Toxicity Test using the Amphipod *Hyalella azteca*," SETAC Platform Session #417, *14th Annual Meeting, Soc. Environ. Toxicol. Chem.*, Nov. 14–18, 1993, Houston, TX (1993).

Kubitz, J.A., E.C. Leweke, J.M. Besser, J.B. Drake III and J.P. Giesy, "Effects of Copper-contaminated Sediments on *Hyalella azteca, Daphnia magna* and *Ceriodaphnia dubia*: Survival, Growth and Enzyme Inhibition," *Arch. Environ. Contam. Toxicol.,* 29: 97-103 (1995).

Kubitz, J.A., J.M. Besser, and J.P. Giesy, "A Two-step Experimental Design for a Sediment Bioassay Using Growth of the Amphipod *Hyalella azteca* for the Test End Point," *Environ. Toxicol. Chem.,* 15: 1783-1792 (1996).

Landrum, R.F. and D. Scavia, "Influence of Sediment on Anthracene Uptake, Depuration, and Biotransformation by the Amphipod *Hyalella azteca*," *Can. J. Fish. Aquat. Sci.,* 40: 298-305 (1983).

Landrum, P.F. and W.R. Faust. 1991. "Effect of Variation in Sediment Composition on the Uptake Rate Coefficient for Selected PCB and PAH Congeners by the Amphipod, *Diporeia sp.*," pp. 263-279, *In:* Aquatic Toxicology and Risk Assessment, vol. 14, M.A. Mayes and M.G. Barron (Eds.) ASTM STP 1124, American Society for Testing and Materials. Philadelphia, PA (1991).

Lenth, R.V., "Statistical Power Calculations," *J. Anim. Sci.* 85(E. Suppl.): E24-E29 (2007).

Leung, J., J.D.S. Witt, W. Norwood and D.G. Dixon. "Implications of Cu and Ni Toxicity in Two Members of the *Hyalella azteca* Cryptic Species Complex: Mortality, Growth, and Bioaccumulation Parameters," *Environ. Toxicol. Chem.,* 35: 2817-2826 (2016).

McGee, B.L., C.E. Schlekat, and E. Reinharz, "Assessing Sublethal Levels of Sediment Contamination Using the Estuarine Amphipod *Leptocheirus plumulosus*," *Environ. Toxicol. Chem.,* 12: 577-587 (1993).

McLeay, D.J., S. Yee, K. Doe, and S. Wade, "Initial Evaluation by EP Laboratories of a 10-day Test for Sediment Toxicity Using the Marine Infaunal Amphipod, *Rhepoxynius abronius*," Unpublished Report Prepared for Environment Canada (EP, C&P) and IGATG by McLeay Associates Ltd., West Vancouver, BC (1989).

McLeay, D., S. Yee, K. Doe, and S. Wade, "Phase-II and Phase-III Studies by Environment Canada Laboratories of 10-day Tests for Sediment Toxicity Using Marine or Estuarine Infaunal Amphipods," Unpublished Report Prepared for Environment Canada (EP, C&P) and IGATG by McLeay Associates Ltd., West Vancouver, BC (1991).

McLeay, D.J., S.G. Yee, K.G Doe, and L.M. Porebski, "Development of 10-day Marine/Estuarine Amphipod Assay for Sediment Toxicity in Support of the Canadian Ocean Dumping Program (CEPA, Part VI)," p. 222-234 in: *Proceedings 18th Annual Aquatic Toxicity Workshop*, September 30–October 3, 1991, Ottawa, ON, Tech. Rep. Fish. Aquat. Sci. No. 1863 (1992).

McLeay, D.J., S. Yee, K. Doe, S. Wade, A. Huybers, and M. Fennell, "Validation of Environment Canada's Ten-day Test for Sediment Toxicity Using Marine or Estuarine Amphipods – Phase VI Study," Unpublished Report Prepared for Environment Canada (EP, C&P) and IGATG by McLeay Associates Ltd., West Vancouver, BC (1993).

Maciorowski, H.D., "Comparison of Lethality of Selected Industrial Effluents Using Various Aquatic Invertebrates Under Laboratory Conditions," *Can. Fish. Mar. Serv. Op. Dir. Tech. Rep. No. CEN/T-75-2* (1975).

Major K., D. J. Soucek, R. Giordano, M. J. Wetzel and F. Soto-Adames, "The Common Ecotoxicology Laboratory Strain of Hyalella azteca is Genetically Distinct From Most Wild Strains Sampled in Eastern North America," *Environ. Toxicol. Chem,.* 32: 2637-2647 (2013).

Maki, A.W., "Modifications of Continuous Flow Test Methods for Small Aquatic Organisms," *Prog. Fish. Cult.,* 39: 172-174 (1977).

MESI (Miller Environmental Sciences Inc.), "Method Development Research for the Improvement of Environment Canada's *Hyalella azteca* Test for Survival and Growth in Sediment," unpublished report prepared by Miller Environmental Sciences Inc. (King City, ON) for Environment Canada's Biological Methods Section (Ottawa, ON), April 8, 2010, (2010).

Milani, D., K.E. Day, D.J. McLeay, and R.S. Kirby, "Recent Intra- and Interlaboratory Studies Related to the Development and Standardization of Environment Canada's Biological Test Methods for Measuring Sediment Toxicity Using Freshwater Amphipods (*Hyalella azteca*) or Midge Larvae (*Chironomus riparius*)," Technical report prepared by National Water Research Institute, Burlington, ON, for the Method Development and Application Section, Environment Canada, Ottawa, ON (1996).

Nebeker, A.V. and C.E. Miller, "Use of the Amphipod Crustacean *Hyalella azteca* in Freshwater and Estuarine Sediment Toxicity Tests," *Environ. Toxicol. Chem.,* 7: 1027-1033 (1988).

Nebeker, A.V., M.A. Cairns, J.H. Gakstatter, K.W. Malueg, G.S. Schuytema, and D.F. Krawczyk, "Biological Methods for Determining Toxicity of Contaminated Freshwater Sediments to Invertebrates," *Environ. Toxicol. Chem.,* 3: 617-630 (1984).

Nebeker, A.V., S.T. Onjukka, M.A. Cairns, and D.F. Krawczyk, "Survival of *Daphnia magna* and *Hyalella azteca* in Cadmium-Spiked Water and Sediment," *Environ. Toxicol. Chem.,* 5: 933-938 (1986).

Nebeker, A.V., G.S. Schuytema, W.L. Griffis, J.A. Barbitta, and L.A. Carey, "Effect of Sediment Organic Carbon on Survival of *Hyalella azteca* Exposed to DDT and Endrin," *Environ. Toxicol. Chem.,* 8: 705-718 (1989).

Nebeker, A.V., S.T. Onjukka, D.G. Stevens, G.A. Chapman, and S.E. Dominguez, "Effects of Low Dissolved Oxygen on Survival, Growth and Reproduction of *Daphnia, Hyalella* and *Gammarus*," *Environ. Toxicol. Chem.,* 11: 373-379 (1992).

Nelson, M.K., P.F. Landrum, G.A. Burton, S.J. Klaine, E.A. Crecelius, T.D. Byl, D.C. Gossiaux, V.N. Tsymbal, L. Cleveland, C.G. Ingersoll, and G. Sasson-Brickson, "Toxicity of Contaminated Sediments in Dilution Series with Control Sediments," *Chemosphere,* 27: 1789-1812 (1994).

Newman, M.C., "What Exactly are you Inferring? A Closer Look at Hypothesis Testing," *Environ. Toxicol. Chem.,* 27: 1013-1126 (2008).

Norberg-King, T., "Development of a Standard Protocol for Testing *Hyalella azteca*," p. 361-386 in: *Proceedings, Tiered Testing Issues for Freshwater and Marine Sediments*, September 16–18, 1992, USEPA, Washington, DC (1992).

Nowierski, M., D.G. Dixon, and U. Borgmann, "Effects of Water Chemistry on the Bioavailability of Metals in Sediment on *Hyalella azteca*: Implications for Sediment Quality Guidelines," *Arch. Environ. Con. Tox.,* 49: 322-332 (2005).

Norwood, W.P., U. Borgmann, and D.G. Dixon, "Interactive Effects of Metal Mixtures on Bioaccumulation in the Amphipod *Hyalella azteca*," *Aquat. Toxicol.,* 84: 255-267 (2007a).

Norwood, W.P., U. Borgmann, and D.G. Dixon, "Chronic Toxicity of Arsenic, Cobalt, Chromium and Manganese to *Hyalella azteca* in Relation to Exposure and Bioaccumulation," *Environ. Pollut*., 147: 262-272 (2007b).

Norwood, W.P., T. Watson-Leung, and D. Milani, "Impacts of Zn-Spiked Sediments on Four Benthic Invertebrates: Implications for Canadian Sediment Quality Guidelines," WSTD Contribution No. 09-048 (2009).

NWRI (National Water Research Institute), "Standard Operating Procedures: Culture and Sediment Bioassay Methods for *Chironomus riparius*, *Hexagenia* spp., *Hyalella azteca*, and *Tubifex tubifex*," K. Day, September 30, 1992 draft document, Rivers Research Branch, Canada Centre for Inland Waters, Environment Canada, Burlington, ON [Unpublished Report] (1992).

OECD (Organisation for Economic Co-operation and Development), "Current Approaches in the Statistical Analysis of Ecotoxicity Data: A Guidance to Application," OECD, Paris (2006).

Othoudt, R.A., J.P. Giesy, K.R. Grzyb, D.A. Verbrugge, R.A. Hoke, J.B. Drake, and D. Anderson, "Evaluation of the Effects of Storage Time on the Toxicity of Sediments," *Chemosphere,* 22: 801-807 (1991).

Paine, M.D. and C.A. McPherson, "Phase IV Studies of 10-day Tests for Sediment Toxicity Using Marine or Estuarine Infaunal Amphipods," Unpublished Report Prepared for Environment Canada (Marine Environment Division, EP, C&P) and IGATG by EVS Consultants Ltd., North Vancouver, BC (1991a).

Paine, M.D. and C.A. McPherson, "Validation of a Ten-day Test for Sediment Toxicity Using Marine or Estuarine Infaunal Amphipods – Phase V Study," Unpublished Report Prepared for Environment Canada (Marine Environment Division, EP, C&P) and IGATG by EVS Consultants Ltd., North Vancouver, BC (1991b).

Pastorok, R.A., D.C. Peek, J.R. Sampson, and M.A. Jacobson, "Ecological Risk Assessment for River Sediments Contaminated by Creosote," *Environ. Toxicol. Chem.,* 13: 1929-1941 (1994).

Phipps, G.L., V.R. Mattson, and G.T. Ankley, "The Relative Sensitivity of Three Freshwater Benthic Macroinvertebrates to Ten Contaminants," *Arch. Environ. Toxicol. Chem.,* 28: 281-286 (1995).

Porebski, L.M. and J.M. Osborne. "The Application of a Tiered Testing Approach to the Management of Dredged Sediments for Disposal at Sea in Canada," *Chem. Ecol.,* 14: 197-214 (1998).

Reynoldson, T.B., K.E. Day, C. Clarke, and D. Milani, "Effect of Indigenous Animals on Chronic End Points in Freshwater Sediment Toxicity Tests," *Environ. Toxicol. Chem.,* 13: 973-977 (1994).

Reynoldson, T.B., R.C. Bailey, K.E. Day, and R.H. Norris, "Biological Guidelines for Freshwater Sediment Based on BEnthic Assessment of SedimenT (the BEAST) Using a Multivariate Approach for Predicting Biological State," *Aust. J. Ecol.,* 20: 198-219 (1995).

Rocchini, R.J., M.J.R. Clark, A.J. Jordan, S. Horvath, D.J. McLeay, J.A. Servizi, A. Sholund, H.J. Singleton, R.G. Watts, and R.H. Young, "*Provincial Guidelines and Laboratory Procedures for Measuring Acute Lethal Toxicity of Liquid Effluents to Fish*," B.C. Ministry of Environment, Victoria, BC (1982).

Sager J.C. and C. McFarlane "Radiation," p. 1-30 in: *Plant Growth Chamber Handbook,* R.W. Langhans and T.W. Tibbits (eds.), North Central Regional Research Publication No. 340, Iowa Agriculture and Home Economics Experiment Station Special Report No. 99, Iowa State University of Science and Technology, Ames Iowa (1997).

Saussure, H., "Memoire Sur Divers Crustaces Nouveaux des Antilles et du Mexique," *Mem. Soc. Phys. Hist. Nat.,* 14(2): 417-496 (1858).

Schlekat, C.E., B.L. McGee, D.M. Boward, E. Reinharz, D.J. Velinsky, and T.L. Wade, "Tidal River Sediments in the Washington, D.C. Area. III. Biological Effects Associated with Sediment Contamination," *Estuaries,* 17: 334-344 (1994).

Schubauer-Berigan, M.K. and G.T. Ankley, "The Contribution of Ammonia, Metals and Nonpolar Organic Compounds to the Toxicity of Sediment Interstitial Water from an Illinois River Tributary," *Environ. Toxicol. Chem.,* 10: 925-939 (1991).

Schubauer-Berigan, M.K., J.R. Dierkes, P.D. Monson, and G.T. Ankley, "pH-dependent Toxicity of Cd, Cu, Ni, Pb and Zn to *Ceriodaphnia dubia, Pimephales promelas, Hyalella azteca* and *Lumbriculus variegatus*," *Environ. Toxicol. Chem.,* 12: 1261-1266 (1993).

Sergy, G., "Recommendations on Aquatic Biological Tests and Procedures for Environmental Protection, C&P, DOE," Environment Canada Report, Environmental Protection, Conservation and Protection, Edmonton, AB [Unpublished Report] (1987).

Sibley, P.K., D.G. Dixon, and D.R. Barton, "Comparison of Pore Water, Bulk Sediment, and Elutriate Toxicity Tests for Determining the Toxicity of Sediments Exposed to BKME to Benthic Macroinvertebrates," SETAC Platform Session #414, *14th Annual Meeting, Soc. Environ. Toxicol. Chem*., Nov. 14–18, 1993, Houston, TX (1993).

Sly, P.G. and W.J. Christie, "Factors Influencing Densities and Distributions of *Pontopoereia hoyi* in Lake Ontario," *Hydrobiologia,* 235/236: 321-352 (1992).

Smith, M.E., L.E. Herrin, and J.M. Lazorchak, "Methods for the Laboratory Culture of the Freshwater Amphipod, *Hyalella azteca*," Unpublished Report, *12th Annual Meeting, Soc. Environ. Toxicol. Chem*., Seattle, WA (1991a).

Smith, M.E., L.E. Herrin, D. M. McMullen, and J.M. Lazorchak, "Test Procedures for the Freshwater Amphipod, *Hyalella azteca*," Unpublished Report, *12th Annual Meeting, Soc. Environ. Toxicol. Chem*., Seattle, WA (1991b).

Smith, M.E., L.E. Herrin, S.L. Brewer, and J.M. Lazorchak, "A Standardized Whole Sediment Exposure for *Hyalella azteca* Testing," SETAC Poster Presentation #TP6B13, *13th Annual Meeting, Soc. Environ. Toxicol. Chem*., Nov. 8–12, 1992, Cincinnati, OH [Unpublished Report] (1992a).

Smith, M.E., L.E. Herrin, S.L. Brewer, and J.M. Lazorchak, "A Standard Water and Control Sediment for Toxicity Testing with *Hyalella azteca*," SETAC Poster Presentation #TP6B12, *13th Annual Meeting, Soc. Environ. Toxicol. Chem*., Nov. 8–12, 1992, Cincinnati, OH [Unpublished Report] (1992b).

Smith, M.E., L. Vigano, J.M. Lazorchak, and A.M. Kneipp, "Sediment Test Methods for a 7-day Fathead Minnow and 7-15-day Rainbow Trout Fish Embryo/Larval," SETAC Platform Session #421, *14th Annual Meeting, Soc. Environ. Toxicol. Chem*., Nov. 14–18, 1993, Houston, TX [Unpublished Report] (1993).

Soucek D.J., A. Dickinson and K.M. Major, "Selection of Food Combinations to Optimize Survival, Growth, and Reproduction of the Amphipod *Hyalella azteca* in Static-Renewal, Water Only Exposures," *Environ. Toxicol. Chem*., 35: 2407-2415 (2016).

Stemmer, B.L., G.A. Burton, Jr., and S. Leibfritz-Frederick, "Effect of sediment spatial variance and collection method on Cladoceran toxicity and indigenous microbial activity determinations," *Environ. Toxicol. Chem.,* 9: 1035-1044 (1990).

Strong, D.R., Jr., "Life History Variation Among Populations of an Amphipod (*Hyalella azteca*)," *Ecology,* 53(6): 1103-1111 (1972).

Suedel, B.C. and J.H. Rodgers Jr., "Responses of *Hyalella azteca* and *Chironomus tentans* to Particle-size Distribution and Organic Matter Content of Formulated and Natural Freshwater Sediments," *Environ. Toxicol. Chem.,* 13: 1639-1648 (1994a).

Suedel, B.C. and J.H. Rodgers Jr., "Development of Formulated Reference Sediments for Freshwater and Estuarine Sediment Testing," *Environ. Toxicol. Chem.,* 13: 1163-1175 (1994b).

Suedel, B.C., J.H. Rodgers Jr., and P.A. Clifford, "Bioavailability of Fluoranthene in Freshwater Sediment Toxicity Tests," *Environ. Toxicol. Chem.,* 12: 155-165 (1993a).

Suedel, B.C., E. Deaver, and J.H. Rodgers Jr., "Reducing Uncertainty in Laboratory Sediment Toxicity Tests," SETAC Platform Session #508, *14th Annual Meeting, Soc. Environ*. *Toxicol. Chem*., Nov. 14–18, 1993, Houston, TX [Unpublished Report] (1993b).

Suedel, B.C., E. Deaver, and J.H. Rodgers, Jr., "Formulated Sediment as a Reference and Dilution Sediment in Definitive Toxicity Tests," *Arch. Envir. Contam. Toxicol.,* 30: 47-52 (1996).

Swartz, R.C., W.A. DeBen, K.A. Sercu, and J.O. Lamberson, "Sediment Toxicity and the Distribution of Amphipods in Commencement Bay, Washington, USA," *Mar. Poll. Bull.,* 13: 359-364 (1982).

Swartz, R.C., D.W. Schults, G.R. Ditsworth, W.A. DeBen, and F.A. Cole, "Sediment Toxicity, Contamination, and Macrobenthic Communities Near a Large Sewage Outfall," p. 152-175 in: *Validation and Predictability of Laboratory Methods for Assessing the Fate and Effects of Contaminants in Aquatic Ecosystems*. T.P. Boyle, (ed). ASTM STP 865, American Society for Testing and Materials, Philadelphia, PA (1985a).

Swartz, R.C., G.R. Ditsworth, D.W. Schults, and J.O. Lamberson, "Sediment Toxicity to a Marine Infaunal Amphipod: Cadmium and its Interaction with Sewage Sludge," *Mar. Environ. Res.,* 18: 133-153 (1985b).

Swartz, R.C., F.A. Cole, D.W. Schults, and W.A. DeBen, "Ecological Changes in the Southern California Bight Near a Large Sewage Outfall: Benthic Conditions in 1980 and 1983," *Marine Ecology,* 31: 1-13 (1986).

Swartz, R.C., P.F. Kemp, D.W. Schults, and J.O. Lamberson, "Effects of Mixtures of Sediment Contaminants on the Marine Infaunal Amphipod, *Rhepoxynius abronius*," *Environ. Toxicol. Chem.,* 7: 1013-1020 (1988).

Swartz, R.C., F.A. Cole, J.O. Lamberson, S.P. Ferraro, D.W. Schults, W.A. DeBen, H. Lee, and R.J. Ozretich, "Sediment Toxicity, Contamination, and Amphipod Abundance at a DDT and Dieldrin-contaminated Site in San Francisco Bay," *Environ. Toxicol. Chem.,* 13: 949-962 (1994).

Taylor L.N. and L. Novak, "Inter-laboratory Validation of Organism Recovery for Use in 42 day Sediment Toxicity Tests With *Hyalella azteca" Environ. Toxicol. Chem*., 36: 1085-1089 (2016).

Taylor L.N., Novak L., Rendas M., Antunes P.M.C., Scroggins R.P., "Validation of a new Standardized Test Method for the Freshwater Amphipod *Hyalella azteca*: Determining the Chronic Effects of Silver in Sediment," *Environ. Toxicol. Chem*., 35: 2430-2438 (2016).

Thursby, G.B., J. Heltshe, K.J. and K.J. Scott, "Revised Approach to Toxicity Test Acceptability Criteria Using a Statistical Performance Assessment," *Environ. Toxicol. Chem.,* 16: 1322-1329 (1997).

Tomasovic, M., F.J. Dwyer, I.E. Greer, and C.G. Ingersoll, "Recovery of Known-age *Hyalella azteca* (Amphipoda) from Sediment Toxicity Tests," *Environ. Toxicol. Chem.,* 14: 1177-1180 (1995).

Trussell, R.P., "The Percent Un-ionized Ammonia in Aqueous Ammonia Solutions at Different pH Levels and Temperatures," *J. Fish. Res. Board Can.,* 29: 1505-1507 (1972).

USEPA (United States Environmental Protection Agency), "Methods for Measuring the Acute Toxicity of Effluents to Freshwater and Marine Organisms," Report EPA/600/4-85/013, W.H. Peltier and C.I. Weber, (eds.), USEPA, Cincinnati,OH (1985a).

USEPA (United States Environmental Protection Agency), "Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms," Report EPA/600/4-85/014, W.B. Horning and C.I. Weber (eds.), 162p., USEPA, Cincinnati, OH (1985b).

USEPA (United States Environmental Protection Agency), "Ambient Water Quality Criteria for Ammonia − 1984," Report EPA 440/5-85-001, USEPA, Washington, DC (1985c).

USEPA (United States Environmental Protection Agency), "Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms," 2nd ed., Report EPA/600/4-89/001, 248 p., USEPA, Cincinnati, OH (1989).

USEPA (United States Environmental Protection Agency), "Field and Laboratory Methods Manual, EMAP Surface Waters, Northeast Lake Pilot," D.J. Klemm and J.M. Lazorchak (eds.), Unpublished Report, Environmental Monitoring Systems Laboratory, USEPA, Cincinnati, OH (1991a).

USEPA (United States Environmental Protection Agency), "Standard Operating Procedure for the Culture of *Hyalella azteca*," J. Denny and S. Collyard, Oct. 25, 1991 draft, 19 p., Environmental Research Laboratory, USEPA, Duluth, MN [Unpublished Report] (1991b).

USEPA (United States Environmental Protection Agency), "Methods for Measuring the Acute Toxicity of Effluents and Receiving Waters to Freshwater and Marine Organisms," Report EPA-600/4-90/027, Office of Research and Development, Environmental Monitoring Systems Laboratory, USEPA, Cincinnati, OH (1991c).

USEPA (United States Environmental Protection Agency),"Methods for Aquatic Toxicity Identification Evaluations. Phase I Toxicity Characterization Procedures. (Second edition)," [Edited by T.J. Norberg-King, D.I. Mount, E.J. Durhan, G.T. Ankley, L.P. Burkhard, J.R. Amato, M.T. Lukasewycz, M.K. Schubauer-Berigan, and L. Anderson-Carnahan] U.S. Environmental Protection Agency, Office of Research and Development, Environmental Research Laboratory, Report EPA/600/6-91/003. National Effluent Toxicity Assessment Center Tech. Rept 18-90, Duluth, MN (1991d).

USEPA (United States Environmental Protection Agency), "Toxicity Identification Evaluation: Characterization of Chronically Toxic Effluents, Phase I," [Prepared by T.J. Norberg-King, D.I. Mount, J.R. Amato, D.A. Jensen and J.A. Thompson]. U.S. Environmental Protection Agency, Office of Research and Devlopment. Report EPA/600/6-91/005F. National Effluent Toxicity Assessment Center Tech. Rept. 02-92, Duluth, MN (1992).

USEPA (United States Environmental Protection Agency), "Methods for Measuring the Acute Toxicity of Effluents and Receiving Waters to Freshwater and Marine Organisms," Report EPA-600/4-90/027F, Office of Research and Development, Environmental Monitoring Systems Laboratory, USEPA, Cincinnati, OH (1993).

USEPA (United States Environmental Protection Agency), "Methods for Measuring the Toxicity and Bioaccumulation of Sediment-associated Contaminants with Freshwater Invertebrates," Report EPA 600/R-94/024, 133 p., USEPA, Duluth, MN (1994a).

USEPA (United States Environmental Protection Agency), "Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Marine and Estuarine Organisms (Second Edition)," Report EPA/600/4-91/003, 483 p., Office of Research and Development, USEPA, Cincinnati, OH (1994b).

USEPA (United.States. Environmental Protection Agency), "1999 Update of Ambient Water Quality Criteria for Ammonia," EPA-823-F-99-013, Office of Water, Washington, DC (1999).

USEPA (United States Environmental Protection Agency), "Methods for Measuring the Toxicity and Bioaccumulation of Sediment-associated Contaminants with Freshwater Invertebrates," 2nd ed. (prepared by the Office of Research and Development, Mid-Continent Ecology Division, USEPA, Duluth, Minnesota, and the Office of Science and Technology, Office of Water, USEPA, Washington, DC), Report EPA 600/R-99/064, 192 p. (2000).

USEPA (United States Environmental Protection Agency), "Methods for Collection, Storage and Manipulation of Sediments for Chemical and Toxicological Analyses: Technical Manual," EPA 823-B-01-002, U.S. Environmental Protection Agency, Office of Water, Washington, DC (2001).

USEPA (United States Environmental Protection Agency), "Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms," 4th ed., Office of Water, Report EPA 821-R-02-013, 350 p., USEPA, Washington, DC (2002).

USEPA (United States Environmental Protection Agency), "Guidance for the Use of Dilution-Water (negative) and Solvent Controls in Statistical Data Analysis for Guideline Aquatic Toxicology Studies," September 25, 2008. Memo from Statistics Workgroup and Aquatic Biology Technical Team to Donald Brady, Director, Environmental Fate and Effects Division. Office of Pesticides Programs, Office of Chemical Safety and Pollution Prevention (OCSPP), Washington, DC (2008).

USEPA (United States Environmental Protection Agency), "National Pollutant Discharge Elimination System Test of Significant Toxicity Implementation Document," Report EPA 833-R-10-003, Office of Wastewater Management, Washington, DC (2010).

USEPA (United States Environmental Protection Agency), "Suggested Requirements and Performance Criteria for Laboratories Conducting Sediment Toxicity Tests," November 23, 2011. Memo from David R. Mount, National health and Environmental Effects Research

Laboratory to D. Scott Ireland, Great Lakes National Program Office (GLNPO), Chicago, IL (2011).

USEPA/USACE (United States Environmental Protection Agency/United States Army Corps of Engineers), "Evaluation of Dredged Material Proposed for Discharge in Inland and Near Coastal Waters – Testing Manual (Draft)," Report EPA-823-B-94-002, prepared by USEPA/USACE, Washington, DC (1994).

Van der Hoeven, N., "Power Analysis for the NOEC: What is the Probability of Detecting Small Toxic Effects on Three Different Species using the Appropriate Standardized Test Protocols?," *Ecotoxicology* 7: 355-361 (1998).

Wang, N., R.J. Erickson, C.G. Ingersoll, C.D. Ivey, E.L. Brunson, T. Augspurger, and M.C. Barnhart, "Influence of pH on the Acute Toxicity of Ammonia to Juvenile Freshwater Mussels (Tafmucket, *Lampsilis siliquoidea*)," *Environ. Toxicol. Chem.,* 27(5): 1141-1146 (2008).

West, C.W., V.R. Mattson, E.N. Leonard, G.L. Phipps, and G.T. Ankley, "Comparison of the Relative Sensitivity of Three Benthic Invertebrates to Copper-contaminated Sediments from the Keweenaw Waterway," *Hydrobiologia,* 262: 57-63 (1993).

Whiteman, F.W., G.T. Ankley, M.D. Kahl, D.M. Rau, and M.D. Balcer, "Evaluation of Interstitial Water as a Route of Exposure for Ammonia in Sediment Tests with Benthic Macroinvertebrates," *Environ. Toxicol. Chem.,* 15: 794-801 (1996).

Williams, D.A., "The Comparison of Several Dose Levels with a Zero Dose Control," *Biometrics,* 28: 519-532 (1972).

Witt, J.D.S., and P.D.N. Hebert, "Cryptic Species Diversity and Evolution in the Amphipod genus *Hyalella* within central glaciated North America: A Molecular Phylogenetic Approach" *Can. J. Fish. Aquat. Sci.* 57: 687-698 (2000).

Yee, S., M. Van Rikxoort, and D. McLeay, "The Effect of Holding Time on *Eohaustorius washingtonianus* During Ten-day Sediment Bioassays and Reference Toxicant Tests," Unpublished Report Prepared for Environment Canada (EP, C&P) and the Inter-Governmental Aquatic Toxicity Group (IGATG), 54 p., North Vancouver, BC (1992).

Zajdlik and Associates Inc., "Statistical Guidance for Ecotoxicological Data Derived Using Environment Canada Standardized Biological Test Methods," unpublished report prepared by Zajdlik and Associates Inc. (Rockwood, ON) for Environment Canada's Biological Methods Section (Ottawa, ON), November (2010).

Zumwalt, D.C., F.J. Dwyer, I.E. Greer, and C.G. Ingersoll, "A Water-renewal System that Accurately Delivers Small Volumes of Water to Exposure Chambers," *Environ. Toxicol. Chem.* 13: 1311-1314 (1994).

Appendix A

Biological Test Methods and Supporting Guidance Documents Published by Environment and Climate Change Canada's Method Development and Applications Unit[a]

Title of Biological Test Method or Guidance Document	Report Number	Publication Date	Applicable Amendments
A. Generic (Universal) Biological Test Methods			
Acute Lethality Test Using Rainbow Trout	EPS 1/RM/9	July 1990	May 1996 and May 2007
Acute Lethality Test Using Threespine Stickleback *(Gasterosteus aculeatus)*	EPS 1/RM/10	July 1990	March 2000
Acute Lethality Test Using *Daphnia* spp.	EPS 1/RM/11	July 1990	May 1996
Test of Reproduction and Survival Using the Cladoceran *Ceriodaphnia dubia*	EPS 1/RM/21 2nd edition	February 2007	—
Test of Larval Growth and Survival Using Fathead Minnows	EPS 1/RM/22 2nd edition	February 2011	—
Toxicity Test Using Luminescent Bacteria *(Photobacterium phosphoreum)*	EPS 1/RM/24	November 1992	—
Growth Inhibition Test Using a Freshwater Alga	EPS 1/RM/25 2nd edition	March 2007	—
Acute Test for Sediment Toxicity Using Marine or Estuarine Amphipods	EPS 1/RM/26	December 1992	October 1998
Fertilization Assay Using Echinoids (Sea Urchins and Sand Dollars)	EPS 1/RM/27 2nd edition	February 2011	—
Toxicity Tests Using Early Life Stages of Salmonid Fish (Rainbow Trout)	EPS 1/RM/28 2nd edition	July 1998	—
Test for Survival and Growth in Sediment Using the Larvae of Freshwater Midges *(Chironomus tentans* or *Chironomus riparius)*	EPS 1/RM/32	December 1997	—

[a] These documents are available for purchase from Publication Catalogue, Environment and Climate Change Canada, Ottawa, Ontario, K1A 0H3, Canada. Printed copies can also be requested by email at: enviroinfo@ec.gc.ca. These documents are freely available in PDF at the following website: www.ec.gc.ca/faunescience-wildlifescience/default.asp?lang=En&n=0BB80E7B-1. For further information or comments, contact the Chief, Biological Assessment and Standardization Section, Environment and Climate Change Canada, Ottawa, Ontario, K1A 0H3.

Title of Biological Test Method or Guidance Document	Report Number	Publication Date	Applicable Amendments
A. Generic (Universal) Biological Test Methods (cont'd.)			
Test for Survival and Growth in Sediment and Water Using the Freshwater Amphipod *Hyalella azteca*	EPS 1/RM/33 2nd edition	January 2013	—
Test for Measuring the Inhibition of Growth Using the Freshwater Macrophyte, *Lemna minor*	EPS 1/RM/37 2nd edition	January 2007	—
Test for Survival and Growth in Sediment Using Spionid Polychaete Worms *(Polydora cornuta)*	EPS 1/RM/41	December 2001	—
Tests for Toxicity of Contaminated Soil to Earthworms *(Eisenia andrei, Eisenia fetida*, or *Lumbricus terrestris)*	EPS 1/RM/43	June 2004	June 2007
Tests for Measuring Emergence and Growth of Terrestrial Plants Exposed to Contaminants in Soil	EPS 1/RM/45	February 2005	June 2007
Test for Measuring Survival and Reproduction of Springtails Exposed to Contaminants in Soil	EPS 1/RM/47	September 2007	—
Test for Growth in Contaminated Soil Using Terrestrial Plants Native to the Boreal Region	EPS 1/RM/56	August 2013	—
B. Reference Methods[b]			
Reference Method for Determining Acute Lethality of Effluents to Rainbow Trout	EPS 1/RM/13 2nd edition	December 2000	May 2007 and February 2016
Reference Method for Determining Acute Lethality of Effluents to *Daphnia magna*	EPS 1/RM/14 2nd edition	December 2000	February 2016
Reference Method for Determining Acute Lethality of Sediment to Marine or Estuarine Amphipods	EPS 1/RM/35	December 1998	—
Reference Method for Determining the Toxicity of Sediment Using Luminescent Bacteria in a Solid-Phase Test	EPS 1/RM/42	April 2002	—
Reference Method for Measuring the Toxicity of Contaminated Sediment to Embryos and Larvae of Echinoids (Sea Urchins or Sand Dollars)	1/RM/58	July 2014	

[b] For this series of documents, a *reference method* is defined as a specific biological test method for performing a toxicity test, i.e., a toxicity test method with an explicit set of test instructions and conditions which are described precisely in a written document. Unlike other generic (multi-purpose or "universal") biological test methods published by Environment and Climate Change Canada, the use of a *reference method* is frequently restricted to testing requirements associated with specific regulations.

Title of Biological Test Method or Guidance Document	Report Number	Publication Date	Applicable Amendments
C. Supporting Guidance Documents			
Guidance Document on Control of Toxicity Test Precision Using Reference Toxicants	EPS 1/RM/12	August 1990	—
Guidance Document on Collection and Preparation of Sediment for Physicochemical Characterization and Biological Testing	EPS 1/RM/29	December 1994	—
Guidance Document on Measurement of Toxicity Test Precision Using Control Sediments Spiked with a Reference Toxicant	EPS 1/RM/30	September 1995	—
Guidance Document on Application and Interpretation of Single-Species Tests in Environmental Toxicology	EPS 1/RM/34	December 1999	—
Guidance Document for Testing the Pathogenicity and Toxicity of New Microbial Substances to Aquatic and Terrestrial Organisms	EPS 1/RM/44 2nd edition	December 2015	—
Guidance Document on Statistical Methods for Environmental Toxicity Tests	EPS 1/RM/46	March 2005	June 2007
Procedure for pH Stabilization During the Testing of Acute Lethality of Wastewater Effluent to Rainbow Trout	EPS 1/RM/50	March 2008	—
Supplementary Background and Guidance for Investigating Acute Lethality of Wastewater Effluent to Rainbow Trout	—	March 2008	—
Guidance Document on the Sampling and Preparation of Contaminated Soil for Use in Biological Testing	EPS 1/RM/53	February 2012	—

Appendix B

Members of the Inter-Governmental Ecotoxicological Testing Group (as of November 2016)

Federal, Environment and Climate Change Canada

Suzanne Agius
Marine Protection Programs Section
Gatineau, Québec

Deborah Austin
Marine Protection Programs Section
Gatineau, Québec

Adrienne Bartlett
Aquatic Ecosystem Protection Research Division
Burlington, Ontario

Lee Beaudette
Science & Technology Laboratories
Ottawa, Ontario

Rene Beaulieu
Prairie & Northern Laboratory for Environmental Testing
Edmonton, Alberta

Christian Blaise (Emeritus)
Centre St. Laurent
Montréal, Québec

Lorraine Brown
Pacific & Yukon Laboratory for Environmental Testing
North Vancouver, British Columbia

Joy Bruno
Pacific & Yukon Laboratory for Environmental Testing
North Vancouver, British Columbia

Julia Brydon
Marine Protection Programs Section
Gatineau, Québec

Craig Buday
Pacific & Yukon Laboratory for Environmental Testing
North Vancouver, British Columbia

Melanie Camplin
Prairie & Northern Laboratory for Environmental Testing
Edmonton, Alberta

Ajith Dias Samarajeewa
Science & Technology Laboratories
Ottawa, Ontario

Heather Dillon
Prairie & Northern Laboratory for Environmental Testing
Edmonton, Alberta

Ken Doe (Emeritus)
Atlantic Laboratory for Environmental Testing
Moncton, New Brunswick

Tamzin El-Fityani
National Guidelines and Standards Office
Ottawa, Ontario

Chris Fraser
Science & Technology Laboratories
Ottawa, Ontario

François Gagné
Fluvial Ecosystem Research
Montréal, Québec

Patricia Gillis
Aquatic Ecosystem Protection Research Division
Burlington, Ontario

Manon Harwood
Québec Laboratory for Environmental Testing
Montréal, Québec

Christina Heise
Prairie & Northern Laboratory for Environmental Testing
Edmonton, Alberta

Ryan Hennessy
Science and Technology Laboratories
Ottawa, Ontario

Natasha Hostal
Prairie & Northern Laboratory for Environmental Testing
Edmonton, Alberta

Dale Hughes
Atlantic Laboratory for Environmental Testing
Moncton, New Brunswick

Paula Jackman
Atlantic Laboratory for Environmental Testing
Moncton, New Brunswick

Stephanie Kvas
Science & Technology Laboratories
Ottawa, Ontario

Heather Lemieux
Science & Technology Laboratories
Ottawa, Ontario

Michelle Linssen-Sauvé
Pacific & Yukon Laboratory for Environmental Testing
Norrth Vancouver, British Columbia

Danielle Milani
Aquatic Ecosystem Impacts Research Division
Burlington, Ontario

Heather Osachoff
Pacific & Yukon Laboratory for Environmental Testing
North Vancouver, British Columbia

Joanne Parrott
Aquatic Ecosystem Protection Research Division
Burlington, Ontario

Linda Porebski
Marine Protection Programs Section
Gatineau, Québec

Juliska Princz
Science & Technology Laboratories
Ottawa, Ontario

Ellyn Ritchie
Science & Technology Laboratories
Ottawa, Ontario

Grant Schroeder
Pacific & Yukon Laboratory for Environmental Testing
North Vancouver, British Columbia

Rick Scroggins
Science & Technology Laboratories
Ottawa, Ontario

Rachel Miliano
Pacific & Yukon Laboratory for Environmental Testing
North Vancouver, British Columbia

Troy Steeves
Atlantic Laboratory for Environmental Testing
Moncton, New Brunswick

David Taillefer
Marine Environmental Protection
Gatineau, Québec

Lisa Taylor
Science & Technology Laboratories
Ottawa, Ontario

Sylvain Trottier
Québec Laboratory for Environmental Testing
Montréal, Québec

Graham van Aggelen
Pacific & Yukon Laboratory for Environmental Testing
North Vancouver, British Columbia

Leana Van der Vliet
Science & Technology Laboratories
Ottawa, Ontario

Brian Walker
Québec Laboratory for Environmental Testing
Montréal, Québec

Peter Wells (Emeritus)
Environmental Conservation Service
Dartmouth, Nova Scotia

Federal, Natural Resources Canada

Melissa Desforges
Mining & Mineral Sciences Laboratory, CANMET
NRCan
Ottawa, Ontario

Morgan King
Mining & Mineral Sciences Laboratory, CANMET
NRCan
Ottawa, Ontario

Philippa Huntsman-Mapila
Mining & Mineral Sciences Laboratory, CANMET
NRCan
Ottawa, Ontario

Carrie Rickwood
Mining & Mineral Sciences Laboratory, CANMET
NRCan
Ottawa, Ontario

Provincial

Richard Chong-Kit
Ontario Ministry of Environment
Etobicoke, Ontario

Olesya Hursky
Saskatchewan Research Council
Saskatoon, Saskatchewan

Lisa Kennedy
Ontario Ministry of the Environment
Etobicoke, Ontario

David Poirier
Ontario Ministry of Environment
Etobicoke, Ontario

Trudy Watson-Leung (Chairperson)
Ontario Ministry of Environment
Etobicoke, Ontario

Private Research Facilities/Others

Eloise Veilleux
Centre d'expertise en analyse environnementale du Québec
Ste. Foy, Québec

Appendix C

Environment and Climate Change Canada, Environmental Protection Service, Regional and Headquarters Offices

Headquarters
351 St. Joseph Boulevard
Place Vincent Massey
Gatineau, Quebec
K1A 0H3

Atlantic Region
45 Alderney Drive
Dartmouth, Nova Scotia
B2Y 2N6

Quebec Region
1550 d'Estimauville Avenue
Québec, Quebec
G1J 0C3

Ontario Region
4905 Dufferin St.
Downsview, Ontario
M3H 5T4

Western and Northern Region
Alberta Office:
4999 – 98th Avenue
Edmonton, Alberta
T6B 2X3

Manitoba Office:
150-123 Main Street
Winnipeg, Manitoba
R3C 4W2

Pacific and Yukon Region
Vancouver Office:
401 Burrard Street
Vancouver, British Columbia
V6C 3S5

Yukon Office:
91782 Alaskan Highway
Yukon
Y1A 5B7

Appendix D

Procedural Variations for Culturing *Hyalella azteca*, as described in Canadian and United States Methodology Documents

Source documents are listed chronologically by originating agency. They can be accessed as:

DFO 1989 represents Borgmann and Munawar (1989) and Borgmann *et al.* (1989). Together, these publications give the culturing and test procedures then in use by the Great Lakes Laboratory for Fisheries and Aquatic Sciences, Department of Fisheries and Oceans (DFO), Burlington, ON.

USFWS 1990 represents Ingersoll and Nelson (1990). This publication gives the culturing and test procedures then in use by the National Fisheries Contaminant Research Center, United States Fish and Wildlife Service, Columbia, MO.

ASTM 1991 a standard guide published by the American Society for Testing and Materials (Philadelphia, PA) for conducting sediment toxicity tests with freshwater invertebrates. This guideline document was published again in 1993. See "ASTM 1991a" and "ASTM 1993" in list of references.

USEPA 1991a includes the draft (April 1991) standard operating procedures for culturing and testing *H. azteca* used by the USEPA Environmental Monitoring Systems Laboratory at Cincinnati, OH, See "USEPA 1991a" in list of references.

USEPA 1991b represents the draft (October 25, 1991) standard operating procedures used for culturing *H. azteca* by the United States Environmental Protection Agency, Environmental Research Laboratory, Duluth, MN, See "USEPA 1991b" in list of references.

USEPA 1991c represents a summary presentation of the culturing procedures investigated by the USEPA Environmental Monitoring Systems Laboratory at Cincinnati, OH. See "Smith *et al.*, 1991a" in list of references.

DFO 1992 is the (unpublished) standard operating procedures used for culturing and testing *H. azteca* by U. Borgmann of the Great Lakes Laboratory for Fisheries and Aquatic Sciences, Department of Fisheries and Oceans (DFO), Burlington, ON. The procedure is the same as that published in Borgmann and Norwood (1993). See "DFO 1992" in list of references.

NWRI 1992 is the (unpublished) standard operating procedures used for culturing and testing *H. azteca* by K. Day of the National Water Research Institute (NWRI), Rivers Research Branch, Canada Centre for Inland Waters, Environment Canada, Burlington, ON. See "NWRI 1992" in list of references.

USEPA 1992 represents Norberg-King (1992), of the USEPA Environmental Research Laboratory at Duluth, MN. Procedures in use at 18 United States and Canadian laboratories for culturing and testing *H. azteca* are summarized.

USFWS 1992 represents Ingersoll (1992), of the United States Fish and Wildlife Service, Columbia, MO. Procedures in use at various laboratories for culturing and testing *H. azteca* are listed.

USEPA 1994a is the published methods for measuring the toxicity and bioaccumulation of sediment associated contaminants with freshwater invertebrates by the United States Environmental Protection Agency (principal authors, C.G. Ingersoll, G.T. Ankley, G.A. Burton, F.J. Dwyer, T.J. Norberg-King, and P.V. Winger). See "USEPA 1994a" in list of references.

1. Source of Brood Stock for Culture

Document[a]	Initial Source
DFO 1989	marshy shoreline of small lake near Burlington, Ontario
USFWS 1990	NI[b]
ASTM 1991	natural freshwater source, another laboratory, or a commercial source
USEPA 1991a	natural freshwater source, another laboratory, or a commercial source
USEPA 1991b	best source from a Lake Superior bay; acceptable sources, other laboratories, commercial suppliers, local collections
USEPA 1991c	USEPA Newtown strain
DFO 1992	marshy shoreline of small lake near Burlington, Ontario
NWRI 1992	CCIW Burlington laboratory (W. Norwood/U. Borgmann)
USEPA 1992	various (St. Louis River, 2 labs;[c] lake near Burlington, 2 labs; Michigan State pond, 1 lab; Nebeker strain, 8 labs; USEPA Newtown, 4 labs)
USFWS 1992	NI (various, depending on the laboratory)
USEPA 1994a	various (avoid wild populations unless the ability of the wild population to cross-breed with existing laboratory populations has been demonstrated)

[a] See preceding page for correct citation.
[b] NI = Not indicated. Source was the USEPA Corvallis strain, as provided by A. Nebeker.
[c] labs = laboratories.

2. Culture Vessels and Loading

Document	Vessel Type	Water Volume (L)	No. of Adult Amphipods/L
DFO 1989	2.5-L pyrex glass jar	1.0	5 to 25
USFWS 1990	80-L glass aquarium	50	NI[a]
ASTM 1991	10-L or 20-L aquarium	NI	NI
USEPA 1991a	8-L aquarium	6 L	NI
USEPA 1991b	2-L battery jar or aquarium	1.0	60
USEPA 1991c	30-mL cup	0.02	100
	1-L glass beaker	NI	80
	8-L aquarium	6.0	17 to 33
	76-L aquarium	40.0	13 to 50
DFO 1992	2.5-L pyrex glass jar	1.0	5 to 25
NWRI 1992	10-L glass aquarium	8.0	20 to 25
	1.2-L glass jar	1.0	20 to 25
USEPA 1992	1-L to 39-L aquarium	0.8 to 38	NI
USFWS 1992	1-L to 100-L[b] aquarium	NI	NI
USEPA 1994a	2-L glass beaker	1.0	50
	2.5-L glass jar	1.0	5 to 25
	80-L aquarium	50.0	NI

[a] NI = Not indicated (depends on method used).
[b] Preferred choice.

3. Water Source, Hardness, and Method of Replacement During Culturing

Document	Water Source	Water Hardness	Method of Replacement
DFO 1989	dechl. tap[a]	130 mg/L	IR[d] (once weekly)
USFWS 1990	well	283 mg/L	FT[e] (~3 times/day)
ASTM 1991	well, surface, dechl. tap,[b] or recon.[c]	optional	IR (25 to 30%/week), or FT (100 mL/min)
USEPA 1991a	as per ASTM 1991	optional	FT (100 mL/min), or IR (≥50%/week)
USEPA 1991b	surface or recon.	NI[f]	IR (once weekly)
USEPA 1991c	well or dilute well[g]	100 mg/L, 200 mg/L	IR (daily) or FT
DFO 1992	dechl. tap	130 mg/L	IR (once weekly)
NWRI 1992	dechl. tap	NI	IR (30%, once weekly)
USEPA 1992	dechl. tap (7), well (4), surface (3), recon. (3)	very soft to very hard	IR or FT
USFWS 1992	various	soft/hard[h]	IR or FT[h]
USEPA 1994a	well, surface, recon.,[i] dechl. tap,[b] estuarine[j]	optional	IR or FT[k]

[a] Dechlorinated municipal tap water.
[b] Dechlorinated water should only be used as a last resort, since dechlorination is often incomplete.
[c] Reconstituted water.
[d] IR = Intermittent renewal.
[e] FT = Flow-through.
[f] NI = Not indicated.
[g] Well water with hardness 200 mg/L diluted to hardness 100 mg/L using deionized water.
[h] Preferred choice.
[i] A recipe is provided for preparing suitable reconstituted water with hardness 90 to 100 mg/L.
[j] *H. azteca* have been cultured in reconstituted salt water with salinities up to 15‰.
[k] Renewal of culture water, with at least one volume addition/d, is recommended. As a minimum, the overlying water volume should be changed at least weekly by siphoning.

4. Temperature, Aeration, and Lighting During Culturing

Document	Water Temp. (°C)	Aeration Conditions	Lighting
DFO 1989	25	none	16L:8D,[a] fluor.,[b] 55 μE/m^2/s[c]
USFWS 1990	20 ± 2	gentle (~2 bubbles/s)	16L:8D, 269 to 538 lux
ASTM 1991	20 ± 2	gentle, if IR[d]	16L:8D, 5382 lux
USEPA 1991a	25 ± 2	gentle, if IR	16L:8D, 5382 lux
USEPA 1991b	25	gentle (air stone)	16L:8D, 1280 lux
USEPA 1991c	25 (FT)[e] 23 (IR)	IR only	16L:8D, 538 to 1076 lux
DFO 1992	25	none	16L:8D, fluor., 55 μE/m^2/s[c]
NWRI 1992	23 ± 1	gentle	16L:8D, 51 μE/m^2/s
USEPA 1992	15 to 25[f]	NI[g]	NI
USFWS 1992	20[h] to 25	moderate	16L:8D, 538 to 1076 lux
USEPA 1994a	23	yes if static or IR	16L:8D, 500 to 1000 lux

[a] Daily photoperiod of 16 hours light and 8 hours dark.
[b] Overhead fluorescent tubes.
[c] In the laboratory used by these investigators, 1 μE/m^2/s = 102.5 lux. Conversion could be different for different types of light.
[d] IR = Intermittent renewal.
[e] FT = Flow-through.
[f] One laboratory at 15°C, three at 20°C, one at 21 ± 2°C, eight at 23°C, four at 25°C.
[g] NI = Not indicated.
[h] Preferred choice.

5. Substrate for Amphipods During Culturing

Document	Description of Substrate Used	Size/Quantity of Substrate
DFO 1989	plastic and cotton gauze	several pieces in jar
USFWS 1990	hard maple leaves previously soaked for 30 d and rinsed for 1 h before use	NI[a]
ASTM 1991	dried maple, alder, birch or poplar leaves, pre-soaked several days and then rinsed	NI
USEPA 1991a	shredded brown paper towel	NI
USEPA 1991b	medicinal gauze sponges, 10 × 10 cm, pre-soaked in culture water for 24 to 48 h	1/jar
USEPA 1991c	single layer of unbleached brown paper towel	NI
DFO 1992	sterile 5 × 10 cm gauze bandage, or 5 × 10 cm piece of 210 µm Nitex™ nylon mesh	1/jar
NWRI 1992	2.5 × 2.5 cm strips of 500 µm Nitex™ nylon mesh, pre-soaked in culture water for 24 h	8/aquarium 1/jar
USEPA 1992	various (gauze, 4 labs;[b] leaves, 4 labs; paper towels, 2 labs; plastic mesh, 2 labs; Nitex™, 1 lab; Nitex™/sand/towels, 1 lab; sediment/towels, 1 lab; plastic/leaves, 1 lab; mesh/towel, 1 lab; none, 1 lab	NI
USFWS 1992	maple leaves,[c] Nitex™ screen, cotton gauze, 3-M base web plastic	NI
USEPA 1994a	various (e.g., cotton gauze, maple leaves, artificial coiled-web material)	NI

[a] NI = Not indicated.
[b] labs = laboratories.
[c] Preferred choice.

6. Feeding During Culturing

Document	Description of Food Used	Quantity per Litre[a]	Feeding Frequency
DFO 1989	TetraMin™ fish food flakes[b]	20 mg	1 to 3×/week
USFWS 1990	hard maple leaves plus ground Tetra™ Standard Mix	NI[c]	*ad libitum*
ASTM 1991	choice of dried maple, alder, birch or poplar leaves; rabbit pellets; ground cereal leaves; fish food pellets; brine shrimp; heat-killed *Daphnia*; green algae and spinach	NI	NI
USEPA 1991a	TetraMin™ fish food flakes + brine shrimp	3.3 mg	1×/day
USEPA 1991b	best success using filamentous algae and YCT;[d] cultured diatoms (*Synedra*) as alternative diet	10 mL YCT algal "pinch"	3×/week 1×/week
USEPA 1991c	ground fish food flakes plus dried algae (*Spirulina* sp.)	50 to 167 mg	2×/day
DFO 1992	TetraMin™ fish food flakes[b]	10 mg	1 to 3×/week
NWRI 1992	Nutrafin™ fish food flakes[b]	2 or 4 drops[e]	2×/week
USEPA 1992	various (single food type, 7 labs; multiple food types, 11 labs)[f]	varied	varied[g]
USFWS 1992	maple leaves,[h] TetraMin™, rabbit chow, diatoms	NI	NI
USEPA 1994a	various (e.g., YCT plus algae; TetraMin™)	varied	varied

[a] Amount of food added per litre of culture water.
[b] Flakes were ground and sifted through a 500 µm mesh nylon screen.
[c] NI = Not indicated.
[d] Yeast, Cerophyll™, and trout chow (USEPA diet for culturing *Ceriodaphnia dubia*).
[e] Two drops of a 100 mg Nutrafin™/mL slurry added per jar; 4 drops per aquarium.
[f] Food types include various rations of yeast, Cerophyll, algae, diatoms, wheat grass, alfalfa, TetraMin™, Nutrafin™, YCT, rabbit pellets, leaves, and paper towels.
[g] For intermittent-renewal cultures, feeding frequencies ranged from 1×/month to 2×/day (47% of labs fed 2×/week); for flow-through cultures, frequencies ranged from 1×/week to 1×/day.
[h] Preferred choice.

7. Harvesting Young for Tests

Document	Description of Procedure	Frequency	No. of Young per Litre[a]
DFO 1989	shake off substrate; filter through 275 µm mesh into petri dish; rinse and sort[b]	once/week	NI[c]
USFWS 1990	rinse portion of mixed-age culture off leaves; filter through 425 µm mesh to obtain animals ≤3 mm; hold overnight in 1-L beaker with aerated water	NI	NI
ASTM 1991	rinse portion of mixed-age culture off leaves; filter through sieves 250 µm (for juveniles) to 425 µm mesh (for adults); hold juveniles ≤24 h in beakers	NI	NI
USEPA 1991a	obtain from adults[d], or sieve daily	daily	NI
USEPA 1991b	pour contents of jar into shallow pan; gently rinse/shake animals off substrate; count and return adults to jar; count young and use or rear for 7 days more	once/week	NI
USEPA 1991c	sieve young released from paired adults	3×/week	33 to 120[e]
DFO 1992	as per Borgmann and Munawar (1989)	once/week	5 to 25
NWRI 1992	pour contents of jar onto 363 µm mesh screen and rinse animals from screen into petri dish; separate young by pipette; count adults; count young and hold	once/week	20 to 35
USEPA 1992	NI	NI	NI
USFWS 1992	NI	NI	NI
USEPA 1994a	various, to obtain 7- to 14-d amphipods	varied	varied

[a] Estimated number of young harvested per litre of culture water.
[b] Separated young kept in jars with 1 L water, 1 piece of 5 × 10 cm gauze and 20 mg TetraMin™ for 2 days before being used in bioassays, to ensure survival and determine numbers of available young.
[c] NI = Not indicated.
[d] Paired adults are placed in 1-L beakers, 25 pairs/beaker, and fed. After 24 h, collect released young.
[e] A 1-L beaker with daily replacement of food and water can yield 120 young/day.

Appendix E

Procedural Variations for Sediment Toxicity Tests Using *Hyalella azteca*, as Described in Canadian and United States Methodology Documents

Source documents are listed chronologically by originating agency. They can be accessed as:

DFO 1989 represents Borgmann and Munawar (1989) and Borgmann *et al.* (1989). Together, these publications give the culturing and test procedures then in use by the Great Lakes Laboratory for Fisheries and Aquatic Sciences, Department of Fisheries and Oceans (DFO), Burlington, ON.

USFWS 1990 represents Ingersoll and Nelson (1990). This publication gives the culturing and test procedures then in use by the National Fisheries Contaminant Research Center, United States Fish and Wildlife Service, Columbia, MO.

ASTM 1991 a standard guide published by the American Society for Testing and Materials (Philadelphia, PA) for conducting sediment toxicity tests with freshwater invertebrates. This guideline document was published again in 1993. See "ASTM 1991a" and "ASTM 1993" in list of references.

USEPA 1991a includes the draft (April 1991) standard operating procedures for culturing and testing *H. azteca* used by the USEPA Environmental Monitoring Systems Laboratory at Cincinnati, OH. See "USEPA 1991a" in list of references.

USEPA 1991b represents a summary presentation of the testing procedures in use at the USEPA Environmental Monitoring Systems Laboratory at Cincinnati, OH. See "Smith *et al*. 1991b" in list of references.

DFO 1992 is the (unpublished) standard operating procedures used for culturing and testing *H. azteca* by U. Borgmann of the Great Lakes Laboratory for Fisheries and Aquatic Sciences, Department of Fisheries and Oceans (DFO), Burlington, ON. The procedure is the same as that published in Borgmann and Norwood (1993). See "DFO 1992" in list of references.

NWRI 1992 is the (unpublished) standard operating procedures used for culturing and testing *H. azteca* by K. Day of the National Water Research Institute (NWRI), Rivers Research Branch, Canada Centre for Inland Waters, Environment Canada, Burlington, ON. See "NWRI 1992" in list of references.

USEPA 1992 represents Norberg-King (1992), of the USEPA Environmental Research Laboratory at Duluth, MN. Procedures in use at 18 United States and Canadian laboratories for culturing and testing *H. azteca* are summarized.

USFWS 1992 represents Ingersoll (1992), of the United States Fish and Wildlife Service, Columbia, MO. Procedures in use at various laboratories for culturing and testing *H. azteca* are listed.

USEPA 1994a is the published methods for measuring the toxicity and bioaccumulation of sediment associated contaminants with freshwater invertebrates by the United States Environmental Protection Agency (principal authors, C.G. Ingersoll, G.T. Ankley, G.A. Burton, F.J. Dwyer, T.J. Norberg-King, and P.V. Winger). See "USEPA 1994a" in list of references.

1. Test Type, Age/Size of Test Animals, Vessel Type, Number of Amphipods per Vessel, Number of Replicates per Treatment, and Test Duration

Document[a]	Test Type	Age/Size of Animal	Test Vessel	No. per Vessel	No. of Replic.	Test Duration (days)
DFO 1989	static[b]	0 to 7 d	2.5-L jar	20	NI[c]	28
USFWS 1990	static	≤3 mm[e]	1-L bkr[g]	20	4	29
	FT[d]	≤3 mm	1-L bkr	20	4	29
ASTM 1991	static	2 to 3 mm[f]	1-L bkr	20	4	≤10 to 30
	static	2 to 3 mm	20-L aquar.	100	≥2	≤10 to 30
	FT	2 to 3 mm	1-L bkr	20	4	≤10 to 30
USEPA 1991a	IR[h]	2 ± 1 d	600 mL	20	4	7
USEPA 1991b	IR	2 ± 1 d	600 mL	20	4	7
DFO 1992	static	0 to 7 d	250-mL bkr	20	4	28
NWRI 1992	static[b]	1 to 10 d	250-mL bkr	15	5	28
USEPA 1992	static[i]	variable[j]	NI	NI	NI	10 to 28[k]
USFWS 1992	FT[l]	7 to 14 d[m]	1 L[n]	20	4 to 5	7 to 28[o]
USEPA 1994a	IR or FT[p]	7 to 14 d	300 mL	10	8[q]	10

[a] See preceding page for correct citation.
[b] Distilled water was added as needed to keep the water level constant.
[c] NI = Not indicated.
[d] FT = Flow-through.
[e] About third instar.
[f] Juvenile animals, second or third instar.
[g] bkr = beaker.
[h] IR = Intermittent renewal.
[i] Ten of 12 laboratories did not replace any water, 2 topped off. Nine of 18 labs also performed tests where water was renewed at frequencies ranging from every 4–6 h to twice per week.
[j] Seven labs, known age; 8 labs, sieve for size/age; 2 labs, mixed age; 1 lab, unknown.
[k] Eight labs, 10 d; 1 lab, 10 to 14 d; 4 labs, 14 d; 1 lab, 20 d; 4 labs, 28 d.
[l] Also static or static-renewal.
[m] Preferred choice, mixed age (~7 to 14 d); also known age (0 to 7 d or 7 to 14 d).
[n] Preferred choice; can range from 25 mL to 100 L.
[o] Preferred choice, 10 days.
[p] Two volume additions/d required, by intermittent (IR) or continuous (FT) replacement.
[q] Depends on test objective. Eight replicates are recommended for routine testing.

2. Test Vessels and Materials

Document	Vessel	Cover	Amount of Sediment	Amount of Water
DFO 1989	2.5-L pyrex screwtop jar	plexiglass sheet	1 to1.5 cm layer	NI[a,b]
USFWS 1990	1-L glass beaker	watch glass	200 mL	800 mL
ASTM 1991	1-L glass beaker 20-L aquarium	watch glass NI	200 mL 2 to 3 cm layer	800 mL 15-cm layer
USEPA 1991a	600 mL	watch glass or glass/plastic sheet	100 mL	400 mL
USEPA 1991b	600 mL	NI	100 mL	400 mL
DFO 1992	250-mL beaker	plastic petri dish[c]	40 mL[d]	160 mL[d]
NWRI 1992	250-mL beaker	petri dish[e]	50 mL	200 mL
USEPA 1992	NI	NI	NI	NI
USFWS 1992	1 L[f]	NI	NI[g]	NI[g]
USEPA 1994a	300-mL high-form lipless beaker	NI	100 mL	175 mL

[a] NI = Not indicated.
[b] Total volume (sediment plus seawater), 1.5 L.
[c] Notch cut out for air supply.
[d] Water added to beaker, then sediment introduced.
[e] Hole drilled for passage of airline tubing.
[f] Preferred choice; can range from 25 mL to 100 L.
[g] Water:sediment ratio can range from 4:1 (preferred choice) to 1:1.

3. Water Source, Hardness, and Method of Replacement During Test

Document	Water Source	Water Hardness	Method of Replacement
DFO 1989	dechlorinated tap[a]	130 mg/L	static with top up[b]
USFWS 1990	reconstituted[c]	134 mg/L	static, FT[d,e]
ASTM 1991	well, surface, dechlorinated tap, or reconstituted	optional	static or flow-through
USEPA 1991a	diluted well[f]	90 to 110 mg/L	IR[g]
USEPA 1991b	diluted well[f]	100 mg/L	IR, daily
DFO 1992	dechlorinated tap[a]	130 mg/L	static[h]
NWRI 1992	dechlorinated tap[a]	NI[i]	static[j]
USEPA 1992	NI	NI	static or renewal[k]
USFWS 1992	NI	soft, hard[l]	static, IR, FT[l]
USEPA 1994a	culture, well, surface, site, or reconstituted[m]	optional	IR or FT, 2×/d[n]

[a] Same source and hardness as used for culturing amphipods.
[b] Distilled water was added as needed to keep water level constant.
[c] Well water with hardness 283 mg/L was used for culturing amphipods; reconstituted water was used during the test.
[d] FT = Flow-through.
[e] 3.8 volume additions per beaker, per day.
[f] Well water with hardness 200 mg/L diluted to hardness 100 mg/L using deionized water.
[g] IR = Intermittent renewal.
[h] Water in controls only should be replaced weekly.
[i] NI = Not indicated.
[j] Water lost by evaporation was replaced weekly using distilled water.
[k] Of 18 laboratories surveyed, 10 used static with no replacement and 2 used static with top up. Nine of the 18 laboratories also renewed overlying water at a frequency ranging from every 4 to 6 h to 2 times/week.
[l] Preferred choice.
[m] A recipe was provided for preparing suitable reconstituted water with hardness 90 to 100 mg $CaCO_3$/L.
[n] Each test vessel should receive two volume additions/d of overlying water, using an intermittent-renewal (manual or automated) or continuous-flow system for replacements.

4. Temperature, Aeration, and Lighting During Test

Document	Water Temp. (°C)	Aeration Conditions	Lighting
DFO 1989	21 ± 1	gentle, using aquarium airstone suspended several cm above sediment	16L:8D,[a] fluor.,[b] 55 µE/m^2/s
USFWS 1990	20 ± 2	gentle (~2 bubbles/s)	16L:8D, 269 to 538 lux
ASTM 1991	20 to 25	gentle	16L:8D, 538 lux
USEPA 1991a	25 ± 1	gentle	16L:8D, 538 lux
USEPA 1991b	25 ± 1	none	NI
DFO 1992	25	gentle, using disposable glass pipette with tip at midpoint of water column	16L:8D, fluor.
NWRI 1992	23 ± 1	gentle	16L:8D
USEPA 1992	20 to 25[c]	NI[d]	NI
USFWS 1992	20 to 25[e]	none or moderate[f]	16L:8D, 269 to 538 lux
USEPA 1994a	23 ± 1[g]	normally, none[h]	16L:8D, ~500 to l000 lux, wide-spectrum fluorescent

[a] Daily photoperiod of 16 hours light and 8 hours dark.
[b] Overhead fluorescent tubes.
[c] Seven laboratories at 20°C, one at 20 to 25°C, four at 23°C, five at 25°C.
[d] NI = Not indicated.
[e] Preferred choice, 20°C.
[f] Preferred choice, none.
[g] Daily mean temperature must be 23 ± 1°C; instantaneous temperature must always be 23 ± 3°C.
[h] Aerate if dissolved oxygen in overlying water drops below 40% of saturation.

5. Feeding During Test

Document	Description of Food Used	Quantity per Vessel	Feeding Frequency
DFO 1989	TetraMin™ fish food flakes[a]	20 mg	1 to 3×/week
USFWS 1990	Purina™ rabbit pellets	14 or 20 mg[b]	3×/week
ASTM 1991	rabbit pellets[c]	varied[b,d]	2 to 3×/week
USEPA 1991a	ground TetraMin™ fish food flakes	14 mg	3×/week
USEPA 1991b	blended fish food flakes	1 mL	Days 0, 2, 4, and 6
DFO 1992	TetraMin™ fish food flakes[a]	5 mg	3×/week
NWRI 1992	Nutrafin™ fish food flakes[a]	8 mg[e]	2×/week
USEPA 1992	NI[f]	NI	varied[g]
USFWS 1992	varied[h]	NI	NI
USEPA 1994a	YCT[i]	1.5 mL	daily

[a] Flakes were ground and sifted through a 500 µm mesh nylon screen.
[b] 14 mg/beaker for static tests; 20 mg/beaker for flow-through tests.
[c] Pellets should be ground, dispersed in deionized water, and resuspended when aliquots are taken.
[d] Options include 6 mg pellets 3×/week for first week, and 12 mg per feeding thereafter.
[e] Added as a slurry of ground Nutrafin™, prepared by adding 1 g flakes to 100 mL distilled water and pulverizing. A volume of ~604 µL is equivalent to 8 mg.
[f] NI = Not indicated.
[g] Of 16 laboratories surveyed, 5 fed 7×/week during tests, 5 fed 3×/week, 2 fed 2×/week, 1 fed 1×/week, 1 fed every 48 h, 1 fed at start only, and 1 did not feed during test.
[h] None; rabbit chow; yeast, Cerophyll™, and trout chow (YCT); maple leaves; or TetraMin™.
[i] Yeast, Cerophyll™, and trout chow.

6. Monitoring Quality of Overlying Water During Test

Document	Variables Monitored[a]	Frequency
DFO 1989	NI[b]	NI
USFWS 1990	DO pH alk hard cond	at least every 10 days, each treatment[c]
ASTM 1991	DO pH alk hard cond temp	beginning, end, and at least weekly[d,e] beginning, end, and at least weekly[d] daily[d,f]
USEPA 1991a	DO pH alk hard cond temp	beginning and end
USEPA 1991b	NI	NI
DFO 1992	DO pH ammonia	beginning and at least weekly[c] beginning (optional but desirable)
NWRI 1992	DO pH cond temp	Days 0, 14, and 28
USEPA 1992	NI	NI
USFWS 1992	NI	NI
USEPA 1994a	DO[g] pH[g] alk hard cond ammonia temp	daily beginning and end[h] daily[d,f]

[a] DO = dissolved oxygen; pH = hydrogen ion concentration; alk = total alkalinity; hard = total hardness; cond = specific conductivity; temp = temperature.

[b] NI = Not indicated.

[c] 50-mL volume of overlying water removed for measurements. In static exposure, this was replaced with fresh, temperature-adjusted overlying water.

[d] Measured in at least one test vessel representing each treatment.

[e] DO to be measured if any interruption of air (static test) or water (flow-through test) and whenever behaviour of animals indicates DO too low (e.g., if amphipods are seen to have emerged from sediment).

[f] Daily mean temperature must be within ±1°C of desired temperature; instantaneous temperature must be within ±3°C of desired temperature.

[g] Can be measured directly, using a probe.

[h] Overlying water should be sampled just before water renewal from about 1 to 2 cm above sediment surface, using a pipet. Values should not vary by more than 50% during a test.

7. Storage and Characterization of Sediment Used in Test

Document	Storage Conditions	Characteristics Measured[a]
DFO 1989	fridge, plastic bags	NI[b]
USFWS 1990	4°C in Teflon™ bags by 24 h; use within 2 weeks	TOC, IC, W, SSC, M, PCB, PAH
ASTM 1991	4 ± 2°C for ≤2 weeks[c]	at least pH, TOC, W, SSC; might include BOD, COD, IC, TVS, AVS, Eh, OG, OS, TA, M, PAH, PW
USEPA 1991a	4°C for ≤2 weeks	NI
USEPA 1991b	NI	NI
DFO 1992	NI	W, TVS, SG
NWRI 1992	sealed plastic buckets, 4°C; use within 6 weeks	NI
USEPA 1992	NI	NI
USFWS 1992	NI	NI
USEPA 1994a	4°C[d]	at least pore water pH + TA, and TOC, W, SSC; might include BOD, COD, CEC, IC, TVS, AVS, Eh, OG, SOC, M, PAH, PW

[a] TOC = total organic carbon; IC = inorganic carbon; W = % water; SSC = % sand, silt, and clay; pH = hydrogen-ion concentration; BOD = biochemical oxygen demand; COD = chemical oxygen demand; CEC = cation exchange capacity; TVS = total volatile solids; SG = specific gravity (g/mL); AVS = acid volatile sulphides; Eh = oxidation reduction potential; OG = oil and grease; OS = organosilicones; TA = total ammonia; M = metals (e.g., As, Cd, Cr, Cu, Hg, Ni, Pb, Zn); PCB = total polychlorinated biphenyls; PAH = polycyclic aromatic hydrocarbons; SOC = synthetic organic compounds; PW = pore water analyses.

[b] NI = Not indicated.

[c] If stored longer than two weeks, sediment should be retested to confirm that toxicity has not changed.

[d] Start test as soon as possible following sample collection. If toxicity test is started after two weeks of collection, it is desirable to conduct additional characterizations of sediment to evaluate possible effects of storage.

8. Manipulation of Sediment Before Use in Test

Document	Sediment Manipulation
DFO 1989	wet-sieved through 275 µm nylon screen, using overlying water in test jars; allowed to settle several days before adding amphipods
USFWS 1990	mixed in storage container; aliquot smoothed with Teflon™ spoon in test beaker; overlying water poured gently along side of beaker; allowed to settle overnight before adding amphipods
ASTM 1991	mix thoroughly; may be wet-press sieved to remove large particles and indigenous organisms; may be diluted and mixed in a 1:1 ratio with overlying water to facilitate sieving; smooth as layer in test vessel; pour overlying water gently along side of beaker; allow to settle overnight before adding amphipods
USEPA 1991a	mix sample; smooth aliquot as layer in test vessel; pour overlying water gently along side of beaker; allow to settle overnight before adding amphipods
USEPA 1991b	NI
DFO 1992	mix sample or take aliquots from several locations; using a stainless steel spoon, transfer aliquot to beaker already containing overlying water; aerate overlying water vigorously for 24 h; reduce aeration to gentle flow and allow sediment to settle for further 24 h
NWRI 1992	wet-sieve through 250 µm mesh using portion of overlying water for test; discard residue retained and leave sieved sediment and overlying water undisturbed overnight; decant overlying water for use in test; add sediment to replicate beakers and then overlying water; allow to settle 24 h and aerate minimum of 1 h before adding amphipods[a]
USEPA 1992	NI[b]
USFWS 1992	NI
USEPA 1994a	mix, including any separated water; samples should not be sieved;[c] remove large organisms and large debris using forceps

[a] In some situations, the sieved sediment is frozen for 24 h and thawed to kill all residual eggs of tubificids.
[b] NI = Not indicated.
[c] If sediment must be sieved, samples should be taken before and after sieving (e.g., pore water metals) to document the influence of sieving on sediment chemistry.

9. Terminating Test and Biological Endpoints

Document	Terminating Test	Biological Endpoints
DFO 1989	sieve contents of jar through 275 µm mesh screen; sort, count, and weigh survivors	mean percent survival mean wet weight
USFWS 1990	wet-sieve sediment; preserve animals in sugar-formalin for subsequent measurement of length[a]	percent survival body length (mm)
ASTM 1991	pipet surviving animals from water column; sieve sediment using 500 µm mesh screen; count live and dead animals; measure length[b]	mean percent survival mean body length mean weight maturation
UEPA 1991a	sieve contents of beaker through 500 µm mesh screen; rinse animals from screen; count live and dead animals; measure mean dry weight[b]	mean percent survival mean dry weight
USEPA 1991b	NI	mean percent survival mean dry weight
DFO 1992	sieve contents of jar through 275 µm mesh screen; sort, count, and weigh survivors	mean percent survival mean wet weight
NWRI 1992	sieve contents of beaker through 500 µm mesh screen; count and weigh surviving amphipods	mean percent survival mean dry weight[c]
USEPA 1992	NI[d]	NI
USFWS 1992	NI	percent survival length or weight maturation
UEPA 1994a	pipet amphipods from water or sediment surface; sieve sediment through 710 µm mesh screen or using multiple sieves; count survivors and measure growth[e]	mean percent survival length or weight[e]

[a] Animals not recovered are presumed to have died and decomposed.
[b] Additional screen sizes may be used for sieving. Animals may be preserved for subsequent determinations of length or weight.
[c] Surviving animals from each beaker are dried for 24 h at 60°C, then weighed as a group.
[d] NI = Not indicated.
[e] Survivors can be preserved in 8% sugar formalin solution for growth (i.e., body length) measurements. If determining dry weight, pool survivors and dry at 60 to 90°C to constant weight, bring to room temperature and weigh to nearest 0.01 mg. Measurement of growth is optional.

10. Use of Control/Reference Sediment and Requirement for Valid Test

Document	Control/Reference Sediment	Requirement for Valid Test
DFO 1989	none used[a]	NI[b]
USFWS 1990	fine-grained control sediment used	NI
ASTM 1991	every test requires a negative (clean) control sediment or a clean reference sediment[c]	mean survival ≥80%[d]
USEPA 1991a	every test requires a negative (clean) control sediment or a clean reference sediment[c]	mean survival ≥80%;[d] single-vessel survival, >70%[d]
USEPA 1991b	negative control sediment used	NI
DFO 1992	none used[a,e]	NI
NWRI 1992	negative control and reference sediments used	mean survival ≥80%[d]
USEPA 1992	negative control sediment used	mean survival 60 to 90%[f]
USFWS 1992	negative control sediment used	mean survival ≥80%[d]
USEPA 1994a	negative control and reference sediments used	mean survival ≥80%[d]

[a] Control survival and growth was measured using gauze as substrate for animals (no sediment).
[b] NI = Not indicated.
[c] A reference sediment should be collected from the field in a clean area, and represent the test sediment in sediment characteristics (e.g., particle size, total organic carbon, pH).
[d] The test is unacceptable if the average survival of organisms in any test vessel containing negative control sediment is less than 80%.
[e] Water in controls (but not the beakers with sediment) was replaced weekly.
[f] Thirteen laboratories used 80% control survival for valid test, two used 70%, one used 90%, and one used 60%.

Appendix F

Procedural Variations for Reference Toxicity Tests Using *Hyalella azteca*, as Described in Canadian and United States Methodology Documents

Source documents are listed chronologically by originating agency. They can be accessed as:

ASTM 1991 a standard guide published by the American Society for Testing and Materials (Philadelphia, PA) for conducting sediment toxicity tests with freshwater invertebrates. This guideline document was published again in 1993. See "ASTM 1991a" and "ASTM 1993" in list of references.

USEPA 1991a includes the draft (April 1991) standard operating procedures for culturing and testing *H. azteca* used by the USEPA Environmental Monitoring Systems Laboratory at Cincinnati, OH. See "USEPA 1991a" in list of references.

USEPA 1991b represents a summary presentation of the testing procedures in use at the USEPA Environmental Monitoring Systems Laboratory at Cincinnati, OH. See "Smith *et al.* 1991b" in list of references.

DFO 1992 is the (unpublished) standard operating procedures used for culturing and testing *H. azteca* by U. Borgmann of the Great Lakes Laboratory for Fisheries and Aquatic Sciences, Department of Fisheries and Oceans (DFO), Burlington, ON. The procedure is the same as that published in Borgmann and Norwood (1993). See "DFO 1992" in list of references.

NWRI 1992 is the (unpublished) standard operating procedures used for culturing and testing *H. azteca* by K. Day of the National Water Research Institute (NWRI), Rivers Research Branch, Canada Centre for Inland Waters, Environment Canada, Burlington, ON. See "NWRI 1992" in list of references.

USEPA 1992a represents Norberg-King (1992), of the USEPA Environmental Research Laboratory at Duluth, Minn. Procedures in use at 18 United States and Canadian laboratories for culturing and testing *H. azteca* are summarized.

USEPA 1992b represents Smith *et al.* (1992a), of the USEPA Environmental Monitoring Systems Laboratory at Cincinnati, OH.

USFWS 1992 represents Ingersoll (1992), of the United States Fish and Wildlife Service, Columbia, MO. Procedures in use at various laboratories for culturing and testing *H. azteca* are listed.

USEPA 1994a is the published methods for measuring the toxicity and bioaccumulation of sediment-associated contaminants with freshwater invertebrates by the United States Environmental Protection Agency (principal authors, C.G. Ingersoll, G.T. Ankley, G.A. Burton, F.J. Dwyer, T.J. Norberg-King, and P.V. Winger). See "USEPA 1994a" in list of references.

1. Reference Toxicant(s), Test Type, Duration, and Frequency of Use

Document[a]	Reference Toxicant(s)	Test Type	Test Duration	Frequency of Use
ASTM 1991	none	none	--	--
USEPA 1991a	$CuSO_4$, KCl, NaCl, Na dodecyl sulphate	water-only[b]	96 h	≥1×/month[c]
USEPA 1991b	NI[e]	IR,[d] water-only	96 h	1×/week
DFO 1992	NI	NI	NI	NI
NWRI 1992	$CuSO_4$	static, water-only	48 h	monthly[f]
USEPA 1992a	variable[g]	NI	NI	NI
USEPA 1992b	KCl	IR, spiked soil[h]	7 days	NI
USFWS 1992	NI	NI	NI	NI
USEPA 1994a	NaCl, KCl, Cd, Cu[i]	static, water-only	96 h	monthly[j]

[a] See preceding page for correct citation.
[b] Amphipods are exposed to a range of concentrations of the reference toxicant dissolved in fresh water; no sediment is present in the test.
[c] If preferred, this test may be performed concurrently with the sediment toxicity tests.
[d] IR = Intermittent renewal, with daily replacement of each test solution.
[e] NI = Not indicated.
[f] Chronic reference toxicant tests with sediment are performed biannually, using $CuSO_4$ and $CdCl_2$.
[g] Of 18 laboratories surveyed, 6 used Cd, 1 used Cr, 3 used Cu, 5 used KCl, 1 used NaCl, and 1 used Zn. Presumably, the remaining 4 laboratories did not use a reference toxicant.
[h] A dry, artificial soil was spiked with serial concentrations of the reference toxicant in solution. Each concentration, which represented the overlying water in the beaker, was replaced daily during the test.
[i] Sodium chloride, potassium chloride, cadmium chloride, and copper sulphate are suitable for use. It might be unrealistic to test more than one or two reference toxicants routinely (i.e., monthly).
[j] Ideally, tests with reference toxicants should be conducted in conjunction with sediment tests.

2. Age/Size of Test Animals, Vessel Type, Volume of Test Material(s), Number of Amphipods per Vessel, and Number of Replicates per Treatment

Document	Age/Size of Animal	Test Vessel	Volume of Solution	Volume of Sediment	No. Per Vessel	No. of Replicates
USEPA 1991a	NI[a]	NI	NI	none	NI	NI
USEPA 1991b	2 ± 1 d	NI	20 mL	none	5	4
NWRI 1992	1 to 10 d	250-mL beaker	200 mL	none	15	3 to 5
USEPA 1992b	3 to 7 d[b]	175-mL beaker	100 mL	25 mL	20	4
USEPA 1994a	7 to 14 d	250-mL beaker[c]	≥100 mL[c]	none	≥10[c]	≥3[c]

[a] NI = Not indicated.
[b] Seven-day-old animals increase the power of the test, due to decreased variability.
[c] Tests can also be conducted using 30-mL plastic cups, 20 mL/cup, 1 amphipod/cup, and ≥10 replicates/cup.

3. Source and Hardness of Water Used in Test, and Variables Monitored

Document	Water Source	Water Hardness	Variables Monitoreda	Monitoring Frequency
USEPA 1991a	NI[b]	NI	NI	NI
USEPA 1991b	diluted well[c]	100 mg/L	NI	NI
NWRI 1992	dechlorinated tap[d]	NI	DO, pH, cond	beginning and end
USEPA 1992b	diluted well[e]	100 mg/L	NI	NI
USEPA 1994a	culture, well, site, surface, reconst.[f]	optional[f]	pH, alk, hard, cond, temp, DO	beginning and end daily

[a] DO = dissolved oxygen; pH = hydrogen ion concentration; cond = specific conductivity; alk = total alkalinity; hard = total hardness; temp = temperature.
[b] NI = Not indicated.
[c] Mixture of well water and deionized water.
[d] Same source and hardness as used for culturing amphipods.
[e] Mixture of well, dechlorinated tap, and deionized water.
[f] A recipe is provided for preparing suitable reconstituted water with hardness 90 to 100 mg $CaCO_3$/L.

4. Temperature, Aeration, and Lighting During Test with Reference Toxicant

Document	Water Temp. (°C)	Aeration Conditions	Lighting
USEPA 1991a	NI[a]	NI	NI
USEPA 1991b	25 ± 1	NI	NI
NWRI 1992	NI	NI	NI
USEPA 1992b	25 ± 1	NI	16L:8D
USEPA 1994a	23	none	16L:8D, ~500 to 1000 lux, wide-spectrum fluorescent

[a] NI = Not indicated.

5. Substrate Used in Test with Reference Toxicant

Document	Description of Substrate Used	Size/Quantity of Substrate
USEPA 1991a	NI[a]	NI
USEPA 1991b	NI	NI
NWRI 1992	2.5 × 2.5 cm strip of 500 µm Nitex™ nylon mesh, presoaked in culture water for 24 h	1/beaker
USEPA 1992b	dry, artificial soil, spiked with serial concentrations of the reference toxicant[b]	25 mL/replicate
USEPA 1994a	Nitex™ screen (110 mesh)	NI

[a] NI = Not indicated.
[b] Each concentration, which represented the overlying water and pore water, was replaced daily.

6. Feeding During Test with Reference Toxicant

Document	Description of Food Used	Quantity per Vessel	Feeding Frequency
USEPA 1991a	NI[a]	NI	NI
USEPA 1991b	*S. capricornutum* and cereal leave extract	0.1 mL algae 0.1 mL cereal	daily
NWRI 1992	NI	NI	NI
USEPA 1992b	algae and Cerophyll™	1 mL	daily
USEPA 1994a	YCT[b]	0.5 mL[c]	Days 0 and 2

[a] NI = Not indicated.
[b] Yeast, Cerophyll™, and trout chow; 1800 mg/L stock.
[c] For 250-mL beaker with 10 amphipods. Use 0.1 mL YCT if 30-mL cup with 1 amphipod.

7. Endpoints and Requirement for Valid Test Using Reference Toxicant

Document	Biological Endpoints	Statistical Endpoints	Requirement for Valid Test
USEPA 1991a	NI[a]	EC50[b]	NI
USEPA 1991b	survival	LC50[c]	NI
NWRI 1992	mean percent survival	EC50	mean survival ≥90%[d]
USEPA 1992b	survival[e]	IC50,[f] IC25,[f] NOEC[g]	NI
USEPA 1994a	survival	LC50	mean survival ≥90%[d]

[a] NI = Not indicated.
[b] EC50 = Median effective concentration.
[c] LC50 = Median lethal concentration.
[d] For controls used in test with reference toxicant.
[e] Mean dry weight was shown to be an insensitive endpoint in tests with KCl.
[f] Inhibiting concentration for a (specified) percent effect (i.e., that causing 50% or 25% inhibition).
[g] NOEC = No-observed-effect concentration.

Appendix G

General Systematics of *Hyalella azteca* (Saussure, 1858)[a]

Taxonomy and Phyletic Relationships

Hyalella azteca (Saussure, 1858) is a member of the talitroidean amphipod family Hyalellidae (Bulycheva, 1957). Superfamily Talitroidea also includes the families Talitridae (beach fleas and sandhoppers), Dogielinotidae (North Pacific sand-burrowers), Najinidae (North Pacific algal borers), and the Hyalidae (kelp grazers). At the time of Bulycheva's revision, the family Hyalellidae contained about 20 described species, all in the genus *Hyalella* and all endemic to South American fresh waters (mostly in Lake Titicaca), except for the North American species, *H. azteca*. The family Hyalellidae was redefined by Bousfield (1979, 1982) to include also the Caribbean coastal marine genus *Parhyalella*, and the antipodean fresh- and brackish-water genera *Chiltonia* (New Zealand), *Austrochiltonia* (Australia), and *Afrochiltonia* (South Africa). These last three genera (of chiltoniins) had been placed in the inquilinous marine family Ceinidae by Barnard (1972) and Barnard and Barnard (1983). Bousfield (1996) has redefined the genus *Hyalella* which now includes about 35 described species (in 4 genera), nearly all confined to South America, but 5 of which (including *H. azteca*) occur in North American fresh waters. The genus is believed to have "split off" from the similar Pacific coastal marine genus *Allorchestes*, probably during the Lower Cretaceous (Gondwana times), when South America was beginning its isolation from the outer southern continental masses (Bousfield, 1984). *H. azteca* is almost certainly a later addition to North American fresh waters into which its immediate ancestors penetrated (northwards from South America) following closure of the Panama isthmus during Pliocene-Miocene epochs.

The basic body parts of *H. azteca* are illustrated in Figure G.1. This species has the typical talitroidean (vs. gammaroidean) features of:

(1) short antenna 1 that lacks an accessory flagellum;

(2) mouthparts in which the mandible has a strong molar but lacks a palp, and the palp of maxilla 1 is vestigial;

(3) gnathopods 1 and 2 that are regularly subchelate and subsimilar in females and immatures, but very unequal in mature males (gnathopod 2 much the larger and more powerful);

[a] Prepared by E.L. Bousfield, Research Associate, Royal British Columbia Museum, 675 Bellevue Street, Victoria, BC V8V 1X4.

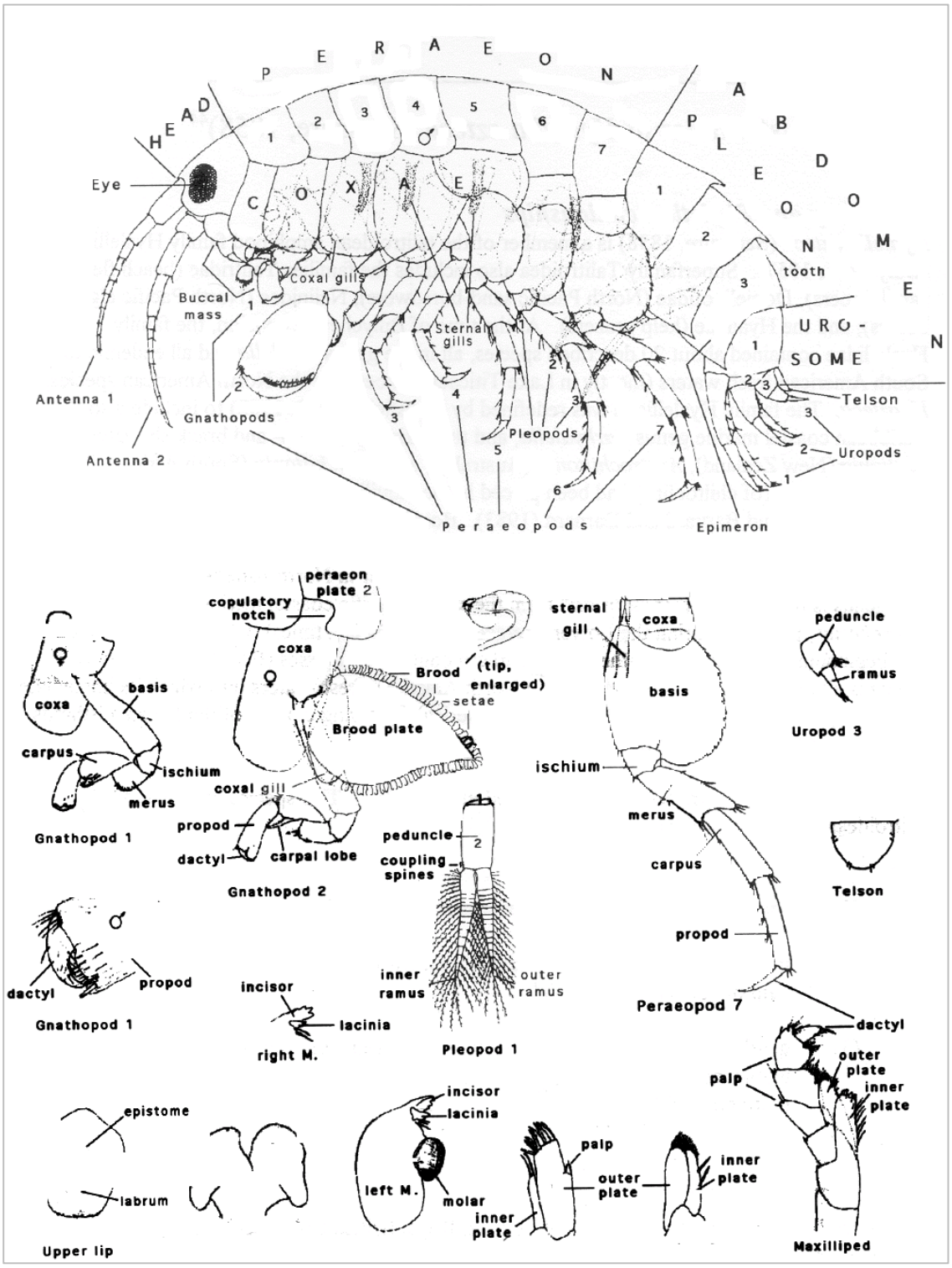


Figure G.1 **Outline of Body, Appendages, and Mouthparts of *Hyalella azteca***
(updated from Bousfield, 1973)

Hyalella azteca (Saussure, 1858)

Figure G.1: Diagnosis (modified from Bousfield, 1973)

Body Length: Male, to 8 mm; female, to 6 mm. Body small, dorsally mucronate on pleon segments 1 and 2, occasionally on 3, or smooth (form *inermis*). Coxal plates very deep, 4th largest; lower margins lightly and evenly spinose. Head, eye subovate, black, slightly larger in male.

Antenna 1, peduncular segments 1 and 2 subequal, flagellum 8–10 segmented. Antenna 2, peduncle slender, segment 5 longer than 4; flagellum 9–10 segmented. Maxilliped, palp segment 2 wider than long, exceeding outer plate.

Gnathopod 1 (male), propod shorter and less deep than carpus, expanding distally; palm oblique, convex. In female, propod narrow, short; palm vertical, convex. Gnathopod 2 (male), propod very large, distally broadest; posterior margin slightly concave; palm convex, with large low tooth near hinge; carpal lobe deep. In female, propod slender, elongate, expanding distally; palm short, convex, vertical.

Peraeopods 3 and 4, posterior margins of segments 5 and 6 with 3–5 short, stout spines. Peraeopods 5–7, basis broadly expanded, posterior margin with 4–10 weak serrations; segments 5 and 6 lacking posterior marginal spines or setae. Abdominal side plates (epimera) 2 and 3, hind corners sharply subquadrate, not produced.

Uropods 1 and 2, both rami with two slender marginal spines. Uropod 3, ramus and peduncle subequal in length, apex with long spine(s). Telson, apex rounded, with two slender wide-set spines.

Coxal gills on peraeopods 2–6 normal, sac-like, smallest on 6. Paired sternal gills at bases of peraeopods 3–5 and 7 are regular in form, not elongate or strongly curved.

With respect to other known North American species of the genus *Hyalella*, *H. azteca* (Saussure) differs in usually possessing a single postero-dorsal tooth or mucronation on each of pleon segments 1 and 2 (occasionally also on 3), and in the relatively elongate, narrow form of the propod and carpus of gnathopod 2 in the female, among other items.

(4) uropod 3 that has only a single short ramus;
(5) telson lobes short or fused (plate-like);
(6) coxal gills located on peraeopods 2-6 only (lacking on peraeopod 7); and
(7) brood plates (female) that are broad, and marginally fringed with short, curl-tipped setae.

H. azteca is a typical member of family Hyalellidae in having the following character states, in combination:

(1) antenna 1 longer than the peduncle of antenna 2;
(2) maxilliped palp with strong dactyl;
(3) gnathopod 2 (female) regular (not mitten-shaped);
(4) lower margin of peraeon 2 (female) with incised "copulatory notch";
(5) peraeopod distal segments not broadened or otherwise modified for burrowing;
(6) pleopods normal (not reduced, modified, or vestigial);
(7) telson entire, plate-like;
(8) coxal gills regular, unmodified; and
(9) sternal gills present variously on inner coxal margin of peraeopods 3–7.

Behaviourally, hyalellids appear to be incapable of jumping in air, as are all members of family Talitridae and many members of family Hyalidae.

With respect to other genera within family Hyalellidae, the genus *Hyalella* differs from *Parhyalella* in possessing sternal gills and a palp on maxilla 1. *Hyalella* differs plesiomorphically from the three austral genera (of chiltoniins, above) in having (in males) sexually mature unmodified pleopods, and a distinct posterior carpal lobe on gnathopod 2.

Distribution and Ecology

According to Bousfield (1958, 1973) and de March (1978), *H. azteca* has been recorded in North America from central Mexico north to about the tree line in Canada and Alaska, and, continent-wide, in virtually all permanent fresh waters that attain a regular summer surface temperature of 10°C or higher. Ecologically, the species prefers fresh waters that are somewhat hard or alkaline, with a normal pH range of 6.0 to 8.0. However, the species has also been found regularly in the upper (tidal) portions of coastal marine estuaries where salinities might reach 2 to 3 ‰ or higher, and in some alkaline lakes where total hardness might exceed 200 mg/L and brine shrimps co-exist (e.g., in some Quill Lakes of Saskatchewan).

With respect to water flow and substrate preferences, the species usually occurs abundantly in lentic waters or ponds where vegetative (especially algal) growth provides food and cover. It is found less frequently in streams and other lotic environments, especially where the bottom consists of uniformly-fine sediments that do not provide protective cover and/or organic food supply.

With respect to respiration, gaseous exchange takes place mainly through the paired coxal gills of peraeopods 2 to 6. Tolerance of low levels of dissolved oxygen and of high levels of carbon dioxide and decomposition gases is apparently higher in this species than in most other North American freshwater amphipods. The paired sternal gills (at the bases of peraeopods 3, 4, 5, and 7) are believed to be mainly osmoregulatory in function, and might facilitate tolerance of a wide range and rapid fluctuation of ionic content within the aquatic medium. Sternal gills might also be partly respiratory in function.

Life Cycle and Reproductive Behaviour

The life cycle of *H. azteca* is essentially annual (Cooper, 1965; Strong, 1972; Conlan and Hendrycks, pers. commun.[b]). In spring, when water temperatures have continuously exceeded 10°C, the overwintering female produces a large clutch of up to 30 eggs. Following hatching and release of juveniles from the brood pouch or marsupium, the female continues to mate and produce further broods. Because of higher ambient summer water temperatures, these later broods occur at more frequent intervals, but clutch sizes tend to be smaller. The newly-hatched juveniles pass through 5 to 6 further instars, or growth stages, before reaching maturity. The spring-spawning females die before onset of the second winter, but late-hatching summer broods comprise the succeeding overwintering population.

With respect to reproductive behaviour, in primitive ("natant" or free-swimming) amphipod superfamilies, pelagic males search out and mate with females freely in the water column, often with a cyclic periodicity. As a member of superfamily Talitroidea, however, *H. azteca* is classified with the reproductively-advanced "reptant" or "bottom-crawling" clade of gammaridean amphipods (Bousfield, 1992; Bousfield and Staude, 1994). In these groups, using a process known as pre-amplexing, precopulation, or mate-guarding (Borowsky, 1984; Conlan, 1990, 1991), males attach themselves dorsally to females, usually by means of their tactile and prehensile first gnathopods. In *Hyalella*, the dactyl of the first gnathopod of the male fits into a special "copulatory notch" (on the lower margin of peraeonal plate 2; see Figure G.1) on both sides of the female's body, leaving the large second gnathopods free to rotate the female into a suitable carrying position and/or to fend off other males. In this "riding" position the male and female remain together, for several hours or days, until the female's next moult. Mating (i.e., amplexus, transfer of sperm) takes place immediately thereafter, on or in bottom substrata, in a short period of time, often within a few seconds, after which the pair separates permanently. During the brief post-exuvial period, the female is especially vulnerable to predation, occasionally by males of competing (or even the same) species, as has been shown in the case of some gammaroideans.

[b] Canadian Museum of Nature (Ottawa, Ontario).

Appendix H

Procedure for Preparing YCT[a] Food for *Hyalella azteca*

Preparing Yeast

1. Add 5.0 g of dry yeast, such as Fleischmann's™, to 1 L of deionized water.

2. Stir with a magnetic stirrer, shake vigorously by hand, or mix with a blender at low speed, until the yeast is well dispersed.

3. Combine the yeast suspension immediately (do not allow to settle) with equal volumes of supernatant from the trout chow and Cerophyll preparations presented subsequently. Discard excess material.

Preparing Cerophyll™ (Dried, Powdered Cereal Leaves)

1. Place 5.0 g of dried, powdered Cerophyll™, cereal leaves, alfalfa leaves, or rabbit pellets[b] in a blender.

2. Add 1 L of deionized water.

3. Mix in a blender at high speed for 5 min, or stir overnight at medium speed on a magnetic stir plate.

4. If a blender is used to suspend the material, place in a refrigerator overnight to settle. If a magnetic stirrer is used, allow to settle for 1 h. Decant the supernatant. Discard excess material.

[a] Mixed ration consisting of yeast, Cerophyll™ (or acceptable substitute), and trout chow (or acceptable substitute). Taken from USEPA (1989) and USEPA (1994a, 2000).

[b] Cerophyll™ can be purchased from Ward's Natural Science Establishment Inc., P.O. Box 92912, Rochester, NY 14692-9012 (716-359-2502). Suitable substitutes for Cerophyll™ include dried, powdered cereal leaves, alfalfa leaves, or rabbit pellets (USEPA, 1994a, 2000). Cereal leaves are available from Sigma Chemical Company, P.O. Box 14508, St. Louis, MO 63178 (800-325-3010). Dried, powdered alfalfa leaves can be obtained from health food stores, and rabbit pellets are available at pet shops.

Preparing Digested Trout Chow[c]

1. Preparation of trout chow requires one week. Use starter or No. 1 pellets[d]

2. Add 5.0 g of trout chow pellets to 1 L of deionized (Milli-Q™ or equivalent) water. Mix well in a blender and pour into a 2-L separatory funnel. Digest before use by aerating continuously from the bottom of the vessel for one week at ambient laboratory temperature. Water lost due to evaporation should be replaced during digestion. Because of the offensive odour usually produced during digestion, the vessel should be placed in a fume hood or other isolated, ventilated area.

3. At the end of the digestion period, place in a refrigerator and allow to settle for a minimum of 1 h. Filter the supernatant through a fine-mesh screen (e.g. Nitex™, 110 mesh). The supernatant can be used fresh, or frozen until use. Discard the remaining particulate material.

Preparing Combined yeast-Cerophyll™-trout chow (YCT) Food

1. Thoroughly mix equal volumes of the three foods as previously described.

2. Place aliquots of the mixture in small (50 to 100 mL) screw-cap plastic bottles.

3. Ideally, food should be stored at 4°C and used within two weeks of preparation. Freshly prepared food can be used immediately, or it can be frozen until needed. Thawed food is stored in the refrigerator between feedings, and is to be used for a maximum of one week after thawed. YCT should not be stored frozen for more three months.

4. Measure the dry weight of solids in each batch of YCT before use. The food should contain 1.7 to 1.9 g solids/L.

[c] USEPA (1994a) indicates that a commercial flaked fish food such as Tetra-min™ may be substituted for trout chow.

[d] Suppliers of trout chow include Zeigler Bros. Inc., P.O. Box 95, Gardners, PA 17324 (phone 717/780-9009); Glencoe Mills, 1011 Elliott, Glencoe, MN 55336 (612/864-3181); and Murray Elevators, 118 West 4800 South, Murray, UT 84107 (800/521-9092).

Appendix I

Logarithmic Series of Concentrations Suitable for Toxicity Tests[a]

Column (Number of concentrations between 10.0 and 1.00, or between 1.00 and 0.10)[b]

1	2	3	4	5	6	7
10.0	10.0	10.0	10.0	10.0	10.0	10.0
3.2	4.6	5.6	6.3	6.8	7.2	7.5
1.00	2.2	3.2	4.0	4.6	5.2	5.6
0.32	1.00	1.8	2.5	3.2	3.7	4.2
0.10	0.46	1.00	1.6	2.2	2.7	3.2
	0.22	0.56	1.00	1.5	1.9	2.4
	0.10	0.32	0.63	1.00	1.4	1.8
		0.18	0.40	0.68	1.00	1.3
		0.10	0.25	0.46	0.72	1.00
			0.16	0.32	0.52	0.75
			0.10	0.22	0.37	0.56
				0.15	0.27	0.42
				0.10	0.19	0.32
					0.14	0.24
					0.10	0.18
						0.13
						0.10

[a] Modified from Rocchini *et al.* (1982).

[b] A series of successive concentrations may be chosen from a column. Midpoints between concentrations in column (x) are found in column (2x + 1). The values listed can represent concentrations expressed as percentage by weight (e.g., mg/kg) or weight-to-volume (e.g., mg/L). As necessary, values can be multiplied or divided by any power of 10. Column 2, which spans two orders of magnitude in concentration, might be used if there was considerable uncertainty about the degree of toxicity. More widely spaced concentrations should not be used, since such usage gives poor resolution of the confidence limits surrounding any threshold-effect value calculated. For effluent testing, there is seldom much gain in precision by selecting concentrations from a column to the right of columns 3 or 4; the finer gradations of columns 4 to 7 might occasionally be useful for testing chemicals that have an abrupt threshold of effect.

Additional information can be obtained at:

Environment and Climate Change Canada
Public Inquiries Centre
12th Floor, Fontaine Building
200 Sacré-Coeur Boulevard
Gatineau QC K1A 0H3
Telephone: 819-938-3860
Toll Free: 1-800-668-6767 (in Canada only)
Email: ec.enviroinfo.ec@canada.ca