



Designation: E1706 – 20

Standard Test Method for Measuring the Toxicity of Sediment-Associated Contaminants with Freshwater Invertebrates¹

This standard is issued under the fixed designation E1706; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reappraisal. A superscript epsilon (ϵ) indicates an editorial change since the last revision or reappraisal.

1. Scope*

1.1 Relevance of Sediment Contamination—Sediment provides habitat for many aquatic organisms and is a major repository for many of the more persistent chemicals that are introduced into surface waters. In the aquatic environment, both organic and inorganic chemicals may accumulate in sediment, which can in turn serve as a source of exposure for organisms living on or in sediment. Contaminated sediments may be directly toxic to aquatic life or can be a source of contaminants for bioaccumulation in the food chain.

1.2 Sediment Assessment Tools—Several types of information may be useful in assessing the risk, or potential risk, posed by sediment contaminants, including: (1) chemical analysis of sediment contaminants; (2) sediment toxicity tests, (3) bioaccumulation tests; and (4) surveys of benthic community structure. Each of these provides a different type of information to the assessment, and integrating information from all four lines of evidence may often provide the most robust assessments.

1.3 Strengths of Toxicity Testing of Contaminated Sediments—Directly assessing the toxicity of contaminated sediments provides some of the same advantages to sediment assessment that whole effluent toxicity testing provides to management of industrial and municipal effluents. As for effluent tests, direct testing of sediment toxicity allows the assessment of biological effects even if: (1) the identities of toxic chemicals present are not (or not completely) known; (2) the influence of site-specific characteristics of sediments on toxicity (bioavailability) is not understood; and (3) the interactive or aggregate effects of mixtures of chemicals present are not known or cannot be adequately predicted. In addition, testing the response of benthic or epibenthic organisms exposed via sediment provides an assessment that is based on the same routes of exposure that would exist in nature, rather than only through water column exposure.

1.4 Relating Sediment Exposure to Toxicity—One of the challenges with sediment assessment is that the toxicity of sediment contaminants can vary greatly with differences in sediment characteristics; a bulk sediment concentration (normalized to dry weight) may be sufficient to cause toxicity in one sediment, while the same concentration in another sediment does not cause toxicity (for example, Adams et al. 1985) (1).² Factors such as the amount and characteristics of the organic carbon present in sediment can alter the bioavailability of many chemicals (Di Toro et al. 1991 (2); Ghosh 2007 (3)), as can other characteristics such as acid volatile sulfide or iron and manganese oxides (Di Toro et al. 1990 (4), Tessier et al. 1996 (5)). Direct measurement of toxicity in contaminated sediments can provide a means to measure the aggregate effects of such factors on the bioavailability of sediment toxicants.

1.5 Understanding the Causes of Sediment Toxicity—While direct testing of sediment toxicity has the advantage of being able to detect the effects of any toxic chemical present, it has the disadvantage of not providing any specific indication of what chemical or chemicals are causing the observed responses. Other techniques, such as spiked-sediment toxicity tests or Toxicity Identification Evaluation (TIE) methods for sediments have been developed and are available to help evaluate cause/effect relationships (USEPA 2007) (6).

1.6 Uses of Sediment Toxicity Tests—Toxicity tests conducted on sediments collected from field locations can be used to: (1) conduct surveys of sediment quality as measured by sediment toxicity; (2) prioritize areas of sediment for more detailed investigation of sediment contamination; (3) determine the spatial extent of sediment toxicity; (4) compare the sensitivity of different organisms to sediment contamination; (5) evaluate the relationship between the degree of sediment contamination and biological effects along a contamination gradient; (6) evaluate the suitability of sediments for removal and placement at other location (for example, dredged material disposal); (7) help establish goals for remedial actions; and (8) assess the effectiveness of remedial actions at reducing sediment toxicity. These applications are generally targeted at

¹ This test method is under the jurisdiction of ASTM Committee E50 on Environmental Assessment, Risk Management and Corrective Action and are the direct responsibility of Subcommittee E50.47 on Biological Effects and Environmental Fate.

Current edition approved April 1, 2020. Published June 2020. Originally approved in 1995. Last previous edition approved in 2010 as E1706 – 05(2010). DOI: 10.1520/E1706-20.

² The boldface numbers in parentheses refer to the list of references at the end of this standard.

*A Summary of Changes section appears at the end of this standard

assessing the likely biological effects of bedded sediments at field sites at the time of sampling. However, toxicity testing of natural or artificial sediments spiked with known quantities of chemicals can also be used to evaluate additional questions such as: (1) determining the potency of a chemical to organisms exposed via sediment; (2) evaluating the effect of sediment composition on chemical bioavailability or toxicity; (3) informing chemical-specific risk assessments for chemicals that may accumulate and persist in sediments upon release; (4) establishing regulatory guidance for chemicals in water or sediment. Spiked sediment studies have the advantage of allowing uni-variate experiments in which exposure gradients can be reliably constructed; as such they lend themselves to the derivation of standardized point estimates of effect, such as a median lethal concentration (LC50) or concentration reducing sublethal performance by a specified amount, such as an effect concentration (for example, EC20 estimated to reduce weight of test organisms by 20 %).

1.7 Limitations—While some safety considerations are included in this standard, it is beyond the scope of this standard to encompass all safety requirements necessary to conduct sediment toxicity tests.

1.8 This standard is arranged as follows:

	Section
Scope	1
Referenced Documents	2
Terminology	3
Summary of Test Methods	4
Significance and Use	5
Interferences	6
Water, Formulated Sediments, Reagents	7
Health, Safety, Waste Management, Biosecurity	8
Facilities, Equipment, and Supplies	9
Sample Collection, Storage, Characterization, and Spiking	10
Quality Assurance and Quality Control	11
Collection, Culturing, and Maintaining the Amphipod <i>Hyalella azteca</i> and the Midge <i>Chironomus dilutus</i>	12
Interpretation of Results and Reporting	13
Precision and Bias	14
Keywords	15
Annexes	
Guidance for 10-d Sediment or Water Toxicity Tests with the Amphipod <i>Hyalella azteca</i>	Annex A1
Guidance for 42-d Sediment or Water Reproductive Toxicity Tests with the Amphipod <i>Hyalella azteca</i>	Annex A2
Guidance for 10-d Sediment or Water Toxicity Tests with the Midge <i>Chironomus dilutus</i>	Annex A3
Guidance for Sediment or Water Life Cycle Toxicity Tests with the Midge <i>Chironomus dilutus</i>	Annex A4
Guidance for Sediment Toxicity Tests with Juvenile Freshwater Mussels	Annex A5
Guidance for Sediment Toxicity Tests with the Midge <i>Chironomus riparius</i>	Annex A6
Guidance for Sediment Toxicity Tests with Mayflies (<i>Hexagenia</i> spp.)	Annex A7
Guidance for Sediment Toxicity Tests with the Oligochaete <i>Tubifex tubifex</i>	Annex A8
References	

1.9 This standard does not purport to address all of the safety concerns, if any, associated with its use. It is the responsibility of the user of this standard to establish appropriate safety, health, and environmental practices and determine the applicability of regulatory limitations prior to use. Specific hazard statements are given in Section 8.

1.10 This international standard was developed in accordance with internationally recognized principles on standard-

ization established in the Decision on Principles for the Development of International Standards, Guides and Recommendations issued by the World Trade Organization Technical Barriers to Trade (TBT) Committee.

2. Referenced Documents

2.1 ASTM Standards:³

- [D1129 Terminology Relating to Water](#)
- [D4387 Guide for Selecting Grab Sampling Devices for Collecting Benthic Macroinvertebrates \(Withdrawn 2003\)⁴](#)
- [E11 Specification for Woven Wire Test Sieve Cloth and Test Sieves](#)
- [E456 Terminology Relating to Quality and Statistics](#)
- [E729 Guide for Conducting Acute Toxicity Tests on Test Materials with Fishes, Macroinvertebrates, and Amphibians](#)
- [E943 Terminology Relating to Biological Effects and Environmental Fate](#)
- [E1241 Guide for Conducting Early Life-Stage Toxicity Tests with Fishes](#)
- [E1325 Terminology Relating to Design of Experiments](#)
- [E1367 Test Method for Measuring the Toxicity of Sediment-Associated Contaminants with Estuarine and Marine Invertebrates](#)
- [E1383 Guide for Conducting Sediment Toxicity Tests with Freshwater Invertebrates \(Withdrawn 1995\)⁴](#)
- [E1391 Guide for Collection, Storage, Characterization, and Manipulation of Sediments for Toxicological Testing and for Selection of Samplers Used to Collect Benthic Invertebrates](#)
- [E1525 Guide for Designing Biological Tests with Sediments](#)
- [E1688 Guide for Determination of the Bioaccumulation of Sediment-Associated Contaminants by Benthic Invertebrates](#)
- [E1733 Guide for Use of Lighting in Laboratory Testing](#)
- [E1847 Practice for Statistical Analysis of Toxicity Tests Conducted Under ASTM Guidelines](#)
- [E1850 Guide for Selection of Resident Species as Test Organisms for Aquatic and Sediment Toxicity Tests](#)
- [E2455 Guide for Conducting Laboratory Toxicity Tests with Freshwater Mussels](#)
- [E3163 Guide for Selection and Application of Analytical Methods and Procedures Used during Sediment Corrective Action](#)
- [IEEE/ASTM-SI-10 Standard for Use of the International System of Units \(SI\): The Modern Metric System](#)

3. Terminology

3.1 The words “must”, “should”, “may”, “can”, and “might” have very specific meanings in this standard. “Must” is used to express an absolute requirement, that is, to state that a test

³ For referenced ASTM standards, visit the ASTM website, www.astm.org, or contact ASTM Customer Service at service@astm.org. For *Annual Book of ASTM Standards* volume information, refer to the standard’s Document Summary page on the ASTM website.

⁴ The last approved version of this historical standard is referenced on www.astm.org.

ought to be designed to satisfy the specified conditions, unless the purpose of the test requires a different design. “Must” is used only in connection with the factors that relate directly to the acceptability of a test. “Should” is used to state that the specified condition is recommended and ought to be met if possible. Although the violation of one “should” is rarely a serious matter, violation of several will often render the results questionable. Terms such as “is desirable,” “is often desirable,” and “might be desirable” are used in connection with less important factors. “May” is used to mean “is (are) allowed to,” “can” is used to mean “is (are) able to,” and “might” is used to mean “could possibly.” Thus, the classic distinction between “may” and “can” is preserved, and “might” is never used as a synonym for either “may” or “can.”

3.2 Definitions—For definitions of other terms used in this test method, refer to Guides E729 and E1241 and Terminology E943, E456, E1325, and D1129. For an explanation of units and symbols, refer to IEEE/ASTM-SI-10.

3.3 Definitions of Terms Specific to This Standard:

3.3.1 clean, *n*—denotes a sediment or water that does not contain concentrations of test materials which cause apparent stress to the test organisms or reduce their survival.

3.3.2 clean sediment and clean water, *n*—denotes a sediment or water that does not contain concentrations of test materials which cause apparent stress to the test organisms or reduce their survival.

3.3.3 concentration, *n*—the ratio of weight or volume of test material(s) to the weight or volume of sediment.

3.3.4 contaminated sediment, *n*—sediment containing chemical substances at concentrations that pose a known or suspected threat to environmental or human health.

3.3.5 control sediment, *n*—a sediment that is essentially free of contaminants and is used routinely to assess the acceptability of a test. Any contaminants in control sediment may originate from the global spread of pollutants and does not reflect any substantial input from local or non-point sources. Comparing test sediments to control sediments is a measure of the toxicity of a test sediment beyond inevitable background contamination. Control sediment is also called a negative control because no toxic effects are anticipated in this treatment.

3.3.6 effect concentration (EC), *n*—the toxicant concentration that would cause an effect in a given percent-age of the test population. Identical to lethal concentration (LC) when the observable adverse effect is death. For example, the EC50 is the concentration of toxicant that would cause a specified effect in 50 % of the test population.

3.3.7 equilibrium partitioning sediment guidelines (ESGs), *n*—numerical concentrations of chemical contaminants in sediment at or below which direct lethal or sublethal toxic effects on benthic organisms are not expected. ESGs are based on the theory that an equilibria exists among contaminant concentration in sediment pore water, contaminant associated with a binding phase in sediment, and biota. ESGs are derived by assigning a protective water-only effects concentration to the pore water (such as a Final Chronic Value), and expressing the

associated equilibrium sediment concentration in terms of the principal binding phase that limits contaminant bioavailability (for example, total organic carbon for nonionic organics or acid volatile sulfides for metals).

3.3.8 formulated sediment, *n*—mixtures of materials used to mimic the physical components of a natural sediment.

3.3.9 inhibition concentration (IC), *n*—the toxicant concentration that would cause a given percent reduction in a non-quantal measurement for the test population. For example, the IC25 is the concentration of toxicant that would cause a 25 % reduction in growth for the test population, and the IC50 is the concentration of toxicant that would cause a 50 % reduction.

3.3.10 interstitial water or pore water, *n*—water occupying space between sediment or soil particles.

3.3.11 lethal concentration (LC), *n*—the toxicant concentration that would cause death in a given percentage of the test population. Identical to EC when the observable adverse effect is death. For example, the LC50 is the concentration of toxicant that would cause death in 50 % of the test population.

3.3.12 lowest-observable-effect concentration (LOEC), *n*—in a toxicity test, the lowest tested concentration of a material at which organisms were adversely affected compared to control organisms as determined by statistical hypothesis tests—should be accompanied by a description of the statistical tests and alternative hypotheses, levels of significance, and measures of performance, for example, survival, growth, reproduction, or development—and must be above any other concentration not producing statistically significant adverse effects.

3.3.13 no-observable-effect concentration (NOEC), *n*—in a toxicity test, the highest tested concentration of a material at which organisms did as well as control organisms as determined by statistical hypothesis tests—should be accompanied by a description of the statistical tests and alternative hypotheses, levels of significance, and measures of performance, for example, survival, growth, reproduction, or development—and must be below any other concentration producing statistically significant adverse effects.

3.3.14 overlying water, *n*—the water placed over sediment in a test chamber during a test.

3.3.15 reference sediment, *n*—a whole sediment near an area of concern used to assess sediment conditions exclusive of material(s) of interest. The reference sediment may be used as an indicator of localized sediment conditions exclusive of the specific pollutant input of concern. Such sediment would be collected near the site of concern and would represent the background conditions resulting from any localized pollutant inputs as well as global pollutant input. This is the manner in which reference sediment is used in dredge material evaluations.

3.3.16 reference-toxicity test, *n*—a test conducted with reagent-grade reference chemical to assess the sensitivity of the test organisms. Deviations outside an established normal range may indicate a change in the sensitivity of the test organism.

population. Reference-toxicity tests are most often performed in the absence of sediment.

3.3.17 *sediment*, *n*—particulate material that usually lies below water. Formulated particulate material that is intended to lie below water in a test.

3.3.18 *spiked sediment*, *n*—a sediment to which a material has been added for experimental purposes.

3.3.19 *whole sediment*, *n*—sediment and associated pore water which have had minimal manipulation. The term bulk sediment has been used synonymously with whole sediment.

4. Summary of Test Method

4.1 *Method Description*—Procedures are described for testing freshwater organisms in the laboratory to evaluate the potential toxicity of chemicals associated with whole sediments or with water-borne exposures to contaminants. Sediments may be collected from the field or spiked with compounds in the laboratory. This standard is a companion to the USEPA (2019) (7) methods manual and both this standard and USEPA (2019) (7) were developed as revisions to the second edition of the USEPA (2000) (8) methods manual and a previous version of this standard (Test Method E1706-19). This standard and USEPA (2019) (7) have lead to the development of other methods for assessing sediment toxicity with invertebrates by other organizations (that is, Environment Canada 1997ab (9, 10), 2007 (11), 2013 (12), 2017 (13); OECD 2004a and b (14, 15), 2006 (16), 2010 (17); ISO 2013 (18); Test Method E1367).

4.2 *Short-term Toxicity Testing with the Amphipod *Hyaella azteca**—Short-term 10-d sediment toxicity testing methods are outlined for the amphipod *Hyaella azteca* in Annex A1. The short-term sediment exposures with *H. azteca* are started with known-age organisms. Toxicity tests are conducted for 10 d in 300-mL chambers containing 100 mL of sediment and 175 mL of overlying water. Overlying water is renewed daily and chemistry of the overlying water is monitored. Food is provided daily. The endpoints in the 10-d toxicity test with *H. azteca* are survival, dry weight, and biomass. Procedures are primarily described for testing freshwater sediments; however, estuarine sediments (up to 15 ‰ salinity) can also be tested in 10-d toxicity tests with *H. azteca*. Also included is guidance on adapting this method for use in testing the toxicity of chemicals introduced via the water column rather than sediment.

4.3 *Long-term Toxicity Testing with the Amphipod *Hyaella azteca**—Methods are described for conducting long-term sediment toxicity tests with *H. azteca* in Annex A2. Toxicity tests are conducted in 300-mL chambers containing 100 mL of sediment and 175 mL of overlying water. Overlying water is renewed daily and chemistry of the overlying water is monitored. Food is provided daily. The long-term sediment exposures with *H. azteca* are started with known-age 7- to 8-d-old amphipods. On Day 28, amphipods are isolated from the sediment and placed in water-only chambers where reproduction is measured on Day 35 and 42. Endpoints measured in the long-term amphipod test include survival (Day 28, 35, and 42), dry weight and biomass (Day 28 and 42), reproduction (number of young per female produced from Day 28 to 42,

number of young/surviving female, and survival-normalized reproduction). Procedures are primarily described for testing freshwater sediments; however, estuarine sediments (up to 15 ‰ salinity) can also be tested in long-term toxicity tests with *H. azteca*. The longer-term method with *H. azteca* also include options for abbreviated versions of this test (for example, 28-d exposures measuring survival, dry weight, and biomass). Also included is guidance on adapting this method for use in testing the toxicity of chemicals introduced via the water column rather than sediment.

4.4 *Short-term Toxicity Testing with the Midge *Chironomus dilutus**—Short-term 10-d sediment toxicity testing methods are outlined the midge *Chironomus dilutus* in Annex A2. The short-term sediment exposures with *C. dilutus* are started with known-age organisms. Toxicity tests are conducted for 10 d in 300-mL chambers containing 100 mL of sediment and 175 mL of overlying water. Overlying water is renewed daily and chemistry of the overlying water is monitored. Food is provided daily. The endpoints in the 10-d toxicity test with *C. dilutus* are survival, ash-free dry weight, and biomass. Also included is guidance on adapting this method for use in testing the toxicity of chemicals introduced via the water column rather than sediment.

4.5 *Long-term Toxicity Testing with the Midge *Chironomus dilutus**—Methods are described for conducting long-term sediment toxicity tests with *C. dilutus* in Annex A4. Midge larvae are exposed to sediments beginning at 3 d old. After 14 d of exposure, a subset of replicates are destructively sampled to determine larval survival, ash-free-dry weight, and biomass. The remaining reproduction replicates are continued through emergence and reproduction of adult midges (for up to about 50 days in exposure started with about 3-d-old larvae), ending when no additional adult emergence has been recorded for 7 consecutive days. Overlying water is renewed daily and chemistry of the overlying water is monitored. Food is provided daily. Endpoints are larval survival, larval weight, larval biomass, percent adult emergence, time to adult emergence, number of egg masses per mated female, average eggs per egg mass, percent of eggs hatching, total young produced, and survival-normalized reproduction. The longer-term method with *C. dilutus* also include options for abbreviated versions of this test (for example, measuring survival, weight, biomass of larvae and emergence of adults but not measuring reproduction). Also included is guidance on adapting this method for use in testing the toxicity of chemicals introduced via the water column rather than sediment.

4.6 *Additional Species for Sediment Toxicity Testing*—Guidance is also provided for conducting sediment toxicity tests with juvenile freshwater mussels (Annex A5), with a second species of midge (*Chironomus riparius*, Annex A6), with a mayfly (*Hexagenia* spp., Annex A7), and with an oligochaete (*Tubifex tubifex*, Annex A8).

4.7 *Bioaccumulation Testing with Sediment*—Guidance for conducting 28-d sediment bioaccumulation tests with the oligochaete *Lumbriculus variegatus* is provided in Guide E1688 and in USEPA (2019) (7).

4.8 The previous version of this standard (Test Method E1706-19) described methods for conducting whole-sediment toxicity tests with *Daphnia magna* and *Ceriodaphnia dubia* (cladocerans) and with *Diporeia* spp. (amphipod). Methods for conducting sediment toxicity tests with cladocerans and with *Diporeia* spp. are not included in the current version of the standard due to limited use of these methods over the past 25 years. A description of the methods for conducting sediment toxicity tests with *D. magna*, *C. dubia* and *Diporeia* spp. can be found in a historic version of the standard (E1706-19) at <http://www.astm.org>.

4.9 Results of toxicity tests, even those with the same species, using procedures different from those described in this standard may not be comparable and using these different procedures may alter bioavailability (Guide E1525). Comparison of results obtained using modified versions of these procedures might provide useful information concerning new concepts and procedures for conducting sediment tests with aquatic organisms. If toxicity tests are conducted with procedures different from those described in this test method, additional tests are required to determine comparability of results. General procedures described in this standard might be useful for conducting tests with other aquatic organisms; however, modifications may be necessary (Guide E1850).

4.10 *Selection of Test Organisms*—A previous version of this standard (Test Method E1706-19), Guide E1525, and USEPA (2000) (8) provide information that was used to select the test organisms in Annex A1 to Annex A8 for conducting sediment toxicity testing.

4.10.1 The choice of a sediment toxicity test organism has a major influence on the relevance, success, and interpretation of a test. Test organism selection should be based on both environmental relevance and practical concerns (Guide E1525, E1850). Ideally, a test organism should: (1) have a toxicological database demonstrating relative sensitivity and discrimination to a range of chemicals of concern in sediment; (2) have a database for inter-laboratory comparisons of procedures (for example, round-robin studies); (3) be in contact with sediment (for example, water column vs. benthic organism); (4) be readily available through culture or from field collection; (5) be easily maintained in the laboratory; (6) be easily identified; (7) be ecologically or economically important; (8) have a broad geographical distribution, be indigenous (either present or historical) to the site being evaluated, or have a niche similar to organisms of concern (for example, similar feeding guild or behavior to the indigenous organisms); (9) be tolerant of a broad range of sediment physico-chemical characteristics (for example, grain size); and (10) be compatible with selected exposure methods and endpoints (Table 1.3 in USEPA 2000) (8). The method should also be (11) peer reviewed (for example, journal articles, USEPA or ASTM methods) and (12) confirmed with responses with natural populations of benthic organisms.

4.10.2 Of these criteria, a database demonstrating relative sensitivity to chemicals, contact with sediment, ease of culture in the laboratory, inter-laboratory comparisons, tolerance to varying sediment physico-chemical characteristics, and confirmation with responses of natural benthic populations were the

primary criteria used for selecting *H. azteca*, *C. dilutus*, *C. riparius*, *Hexagenia* spp., *T. tubifex*, and freshwater mussels for sediment toxicity testing (USEPA 2000 (8), Test Method E1706-19, Guide E1525, Guide E2455).

5. Significance and Use

5.1 Sediment provides habitat for many aquatic organisms and is a major repository for many of the more persistent chemicals that are introduced into surface waters. In the aquatic environment, most anthropogenic chemicals and waste materials including toxic organic and inorganic chemicals can accumulate in sediment, which can in turn serve as a source of exposure for organisms living on or in sediment. Contaminated sediments may be directly toxic to aquatic life or can be a source of contaminants for bioaccumulation in the food chain.

5.2 The objective of a sediment test is to determine whether chemicals in sediment are harmful to or are bioaccumulated by benthic organisms. The tests can be used to measure interactive toxic effects of complex chemical mixtures in sediment. Furthermore, knowledge of specific pathways of interactions among sediments and test organisms is not necessary to conduct the tests. Sediment tests can be used to: (1) determine the relationship between toxic effects and bioavailability, (2) investigate interactions among chemicals, (3) compare the sensitivities of different organisms, (4) determine spatial and temporal distribution of contamination, (5) evaluate hazards of dredged material, (6) measure toxicity as part of product licensing or safety testing, (7) rank areas for clean up, and (8) estimate the effectiveness of remediation or management practices.

5.3 Results of toxicity tests on sediments spiked at different concentrations of chemicals can be used to establish cause and effect relationships between chemicals and biological responses. Results of toxicity tests with test materials spiked into sediments at different concentrations may be reported in terms of a LC50 (median lethal concentration), an EC50 (median effect concentration), an IC50 (inhibition concentration), or as a NOEC (no observed effect concentration) or LOEC (lowest observed effect concentration). However, spiked sediment may not be representative of chemicals associated with sediment in the field. Mixing time, aging and the chemical form of the material can affect responses of test organisms in spiked sediment tests (10.6).

5.4 Evaluating effect concentrations for chemicals in sediment requires knowledge of factors controlling their bioavailability. Similar concentrations of a chemical in units of mass of chemical per mass of sediment dry weight often exhibit a range in toxicity in different sediments (Di Toro et al. 1990 (4), 1991 (2)). Effect concentrations of chemicals in sediment have been correlated to interstitial water concentrations, and effect concentrations in interstitial water are often similar to effect concentrations in water-only exposures. The bioavailability of nonionic organic compounds and metals in sediment is often inversely correlated with the organic carbon concentration; moreover, the bioavailability of metals in sediment are often inversely correlated with acid volatile sulfide. Whatever the route of exposure, these correlations of effect concentrations to

interstitial water concentrations indicate that predicted or measured concentrations in interstitial water can be used to quantify the exposure concentration to an organism. Therefore, information on partitioning of chemicals between solid and liquid phases of sediment is useful for establishing effect concentrations (DiToro et al. 1990 (4), 1991 (2); Wenning et al. 2005 (19)).

5.5 Field surveys can be designed to provide either a qualitative reconnaissance of the distribution of sediment contamination or a quantitative statistical comparison of contamination among sites. Surveys of sediment toxicity are usually part of more comprehensive analyses of biological, chemical, geological, and hydrographic data (USEPA 2002a, b, and c) (20-22). Statistical correlations may be improved and sampling costs may be reduced if subsamples are taken simultaneously for sediment tests, chemical analyses, and benthic community structure.

5.6 Table 1 lists several approaches used to assess of sediment quality. These approaches include: (1) equilibrium partitioning sediment guidelines (ESGs; USEPA 2003 (23), 2005 (24); Nowell et al. 2016 (25)), (2) empirical sediment quality guidelines (for example, probable effect concentrations, PECs; MacDonald et al. 2000 (26), Ingersoll et al. 2001 (27)), (3) tissue residues, (4) interstitial water toxicity, (5) whole-

sediment toxicity with field-collected sediment tests and with sediment-spiking tests, (6) benthic community structure, and (7) sediment quality triad integrating data from sediment chemistry, sediment toxicity and benthic community structure (Burton 1991 (28), Chapman et al. 1997 (29), USEPA 2002a, b, and c (20-22)). The sediment assessment approaches listed in Table 1 can be classified as numeric (for example, ESGs), descriptive (for example, whole-sediment toxicity tests), or a combination of numeric and descriptive approaches (for example, PECs). Numeric methods can be used to derive chemical-specific effects-based sediment quality guidelines (SQGs). Although each approach can be used to make site-specific decisions, no one single approach can adequately address sediment quality. Overall, an integration of several methods using the weight of evidence is the most desirable approach for assessing the effects of contaminants associated with sediment (USEPA 2002a, b, and c (20-22), Wenning et al. 2005 (19), Guide E1525, Guide E3163). Hazard evaluations integrating data from laboratory exposures, chemical analyses, and benthic community assessments (the sediment quality triad) provide strong complementary evidence of the degree of pollution-induced degradation in aquatic communities (Burton 1991 (28), Chapman et al. 1997 (29)). Importantly, the weight

TABLE 1 Sediment Quality Assessment Procedures (Modified from USEPA 1992 (30))

Method	Type			Approach
	Numeric	Descriptive	Combination	
Equilibrium Partitioning Sediment Guidelines (ESGs)		*		An ESG for a given contaminant is determined by calculating the sediment concentration of the contaminant that corresponds to an interstitial water concentration equivalent to the USEPA water-quality criterion for the contaminant.
Empirical Sediment Quality Guidelines	*	*	*	The sediment concentration of contaminants associated with toxic responses measured in laboratory exposures or field assessments (that is, Apparent Effects Threshold (AET), Effect Range Median (ERM), Probable Effect Level (PEL), Probable Effect Concentration (PEC)).
Tissue Residues	*			Safe sediment concentrations of specific chemicals are established by determining the sediment chemical concentration that results in acceptable tissue residues.
Interstitial-water Toxicity	*	*	*	Toxicity of interstitial water isolated from sediment is quantified and identification evaluation procedures are applied to identify and quantify chemical components responsible for sediment toxicity.
Whole-sediment Toxicity with Field-collected Sediments and with Sediment Spiking	*	*	*	Test organisms are exposed to whole sediments that may contain known or unknown quantities of potentially toxic chemicals. Dose-response relationships can be established by exposing test organisms to whole sediments that have been spiked with known amounts of chemicals or mixtures of chemicals.
Benthic Community Structure		*		Environmental degradation is measured by evaluating alterations in resident benthic community structure.
Sediment Quality Triad	*	*	*	Sediment chemical contamination, sediment toxicity, and benthic community structure are measured on the same sediment sample from the site of interest. Correspondence between sediment chemistry, toxicity, and field effects is used to determine sediment concentrations that discriminate conditions of minimal, uncertain, and major biological effects.

of the evidence needed to make a decision (number of methods used) should be determined based on the weight (cost) of the decision.

6. Interferences

6.1 General:

6.1.1 Interferences in the Context of Sediment Toxicity Testing—In narrow terms, the purpose of a sediment toxicity test is to determine whether the constituents of a sediment sample reduce performance of the test organism relative to a control sediment, reference sediment, or some other sediment whose characteristics serve as a meaningful point of comparison. Sediments that reduce relative organism performance are generally considered to be “toxic”, or at least to have greater toxicity than the sediment serving as a point of reference. Because the methods in this standard are intended to assess sediment-associated contaminants, reduced organism performance is generally attributed to the presence of those contaminants. So an interference in a sediment toxicity test can be thought of as a factor that causes a sediment to be judged non-toxic (for example, not different from control) when in fact the level of sediment contamination is sufficient that it should decrease performance (a false negative) or a factor that causes a sediment to indicate toxicity (for example, lower performance relative to control) when in fact the reduced performance is not caused by sediment contaminants. In cases where a gradient in response is being assessed rather than simply “toxic” or “not toxic”, these effects could be viewed as causing greater or lesser response than would be expected absent the influence of the “interfering” factor.

6.1.2 Differences in Responses to Field-collected Sediments when Tested in the Laboratory—Many applications of sediment toxicity tests involve collection of bedded sediments from the field, that are subsequently tested in the laboratory. It is possible that differences between the physical or biological setting of the original field sediment and the conditions of a laboratory sediment toxicity test could create differences in the apparent adverse effects that might result from exposure; this is possible not only because of differences in the conditions themselves, but also because of changes to the sediment that result from its removal from the field, storage, and manipulation as part of the preparation for and conducting laboratory of sediment toxicity testing. Whether this potential is viewed as an “interference” or simply a consequence of the measurement depends on the presumption of the investigator. Laboratory sediment toxicity tests are often used as tools to assess the likely effects of sediment contaminants under field conditions, but this connection is not intrinsic to the test itself. The extrapolation of responses measured in the laboratory to those that might exist in the field is an important, but separate evaluation that is the responsibility of those designing and implementing the overall sediment assessment program. It is worth noting that the issue of laboratory to field extrapolation is by no means unique to contaminated sediment assessment, and while the issue is important to consider, the precedent of using laboratory tests to develop assessment guidance for pollution in natural systems is extensive, such as in the development of water quality guidelines from single species

aquatic toxicity tests (for example, USEPA water quality criteria; Stephan et al. 1985 (31)).

6.1.3 Scope of Interferences Discussion—Because the definition of an interference in a sediment toxicity test is somewhat context sensitive, the remainder of this section does not attempt to define issues specifically as to whether these factors should be considered interferences. Instead, several factors are discussed that are known or suspected to be potential influences on the responses (including chemical accumulation) of organisms exposed as described in this standard. The importance and implications of these factors for specific studies is left to the investigator or the authority under which the study is conducted.

6.2 Issues to Consider in Planning or Evaluating Sediment Tests:

6.2.1 Studies with Spiked Sediments versus Field-collected Sediments—Typically, spiked sediment tests are structured so that the same sediment is tested with differing levels of chemical added. This provides much greater consistency in test conditions across treatments than may exist in studies of field samples collected at multiple sites, which may differ in many different characteristics (for example, grain size, organic carbon) beyond just contaminant concentration(s). As such, most of the issues discussed in 6.2.2 through 6.2.6 are likely to be of greater concern in studies involving multiple samples collected from field sites.

6.2.2 Sediment Collection and Handling Procedures—The processes involved in removing a bedded field sediment from its field location site, and transporting, storing, and preparing the sediment sample for testing have the potential to alter the characteristics of the sediment sample relative to conditions occurring in the field. Section 10 describes a number of steps that can be taken to minimize undesirable changes associated with these processes. Assessment of sediment chemistry at appropriate points in time from collection to completion of testing can provide important information on the degree to which handling, processing, and testing have affected test sediments.

6.2.3 Grain Size—The organisms used in sediment tests described in this standard were selected in part to be tolerant of a fairly wide range of grain size (see Section 1 in USEPA 2000 (8)). Because the interactions of sediment contaminants with sediment particles are generally surface-mediated processes, large particles (for example, <5 mm) with small surface area to volume ratios may have limited influence on the contaminant-related response of organisms, and coarse sieving to remove these large particles is discussed in Section 10 and in Annex A1 to Annex A8. If there is a reason for concern that grain size may be an important influence on study results, testing of field reference samples, or control sediments amended to adjust particle size, may be a useful addition to a study.

6.2.4 Organic Carbon—Organic carbon content of sediments is known to affect the bioavailability of many sediment-associated contaminants, both by the relative amount of organic carbon present (Di Toro et al 1991 (2), 2005 (32)), or by the nature and source of the particles comprising the organic carbon fraction (Ghosh 2007) (3). While the influence of organic

of organic carbon can affect the relationship between contaminant concentration and effect, this would not generally be considered an interference. Several studies have evaluated the effect of different types, concentrations, or sources of organic carbon on the survival and growth of *H. azteca* and *C. dilutus* (also *C. riparius*), and come to varying conclusions depending in part on how the effect is assessed (for example, Lacey et al. 1999 (33), Ristola et al. 1999 (34), Suedel and Rogers 1994 (35), Ankley et al. 1994a (36)). Some of these studies did not include supplemental feeding, and all were conducted before the development of the feeding regimes included in the current sediment toxicity test methods described in this standard. This is important because in studies with no feeding, or insufficient feeding, responses to varying organic carbon can reflect the potential for organic carbon to be a food source, in addition to any other influence on organism performance. In an inter-laboratory study with *H. azteca* (Ivey et al. 2016) (37) which used the currently recommended feeding regimes described in Table A1.1 and Table A2.1, organisms provided with only silica sand as a substrate generally showed performance at or near that achieved using a field control sediment, demonstrating that there is no minimum organic carbon content for the substrate required for *H. azteca* to meet minimum control performance requirements (Tables A1.1 and A1.2). Likewise, *C. dilutus* fed as specified in Annex A3 and Annex A4 readily meet the control performance criteria associated with those tests (USEPA-Duluth, USGS-Columbia, unpublished data). Ankley et al. (1994a) (36) did not find a relationship between sediment organic content (up to 8 %) and 10-d growth of *H. azteca* or *C. dilutus* in a survey of 50 sediments. However, Sampson et al. (2009) (38) reported that very high organic carbon (for example, >20 %) was associated with decreased growth of *C. dilutus*. As for any sediment characteristic that lies outside a typical range, inclusion of control or reference sediments that represent extremes in sediment characteristics may be a useful addition to a study.

6.2.5 Nutritional Content of Sediment—For studies of field-collected sediments, the innate nutritional content of sediments may vary across sediments, which might influence organism performance, especially growth. There has been research evaluating the relationship between different measures of sediment chemistry and growth of benthic organisms (for example, Vos et al. 2000 (39), 2002 (40)). In studies supporting the development of the revised foods and rations described in this standard, a major emphasis was placed on finding feeding regimes that were sufficient to make organism performance largely independent of the presence of additional food sources within a sediment, and thereby reduce any confounding of the response to sediment contaminants. Particularly for *C. dilutus*, increasing rations beyond what is recommended in this standard resulted in even higher larval weights, but larger rations also caused large reductions in dissolved oxygen, and the recommended ration was developed as a compromise between increasing growth and limiting effects on dissolved oxygen. Accordingly, while improvements were made, there remains some potential for organism performance to be enhanced in sediments exceptionally rich in nutritional content. The significance of this as a confounding factor in assessing toxicity from

sediment-associated contaminants is not clear; while clean sediments that have high nutritional content could result in growth above that found in control or reference treatments, it is not clear how much toxicant stress can be offset by the greater nutritional resources, thus potentially interfering with the ability of the test to detect toxicant stress. In the reverse case, where test sediments might have lower nutritional content than the control or reference, a non-nutritive substrate, such as a sand control, can be used as a point of comparison for organism performance when organisms have only the nutrition provided to all sediments. A sand control provides a reference for the low end of performance that might be expected based on limited nutritional content, but it is not necessary that a field sediment perform more poorly than a sand control to be considered potentially toxic.

6.2.6 Density Dependent Growth—As discussed in 6.2.5, the ration provided to *C. dilutus* does not support maximum growth rates. In cases where one or more midge larvae do not survive, there will be proportionately more food available to the survivors, raising the potential for higher growth. This potential was supported by a meta-analysis of growth and survival in control replicates from a large number of 10-d sediment tests with *C. dilutus*, which showed a general tendency toward higher than average weights in replicates with lower than average survival, and vice versa (Fig. A3.2; USEPA-Duluth unpublished data). The effect of density-dependent growth can be compensated for, to some degree, by analyzing total biomass in addition to average dry weight. In the same meta-analysis, control biomass was found to have a lower coefficient of variation than did average dry weight. A similar analysis of data from 10-d tests with *H. azteca* did not show a relationship between survival (density) and weight, indicating that this issue is of less concern for *H. azteca* (Fig. A1.3).

6.2.7 Ammonia—Ammonia is produced naturally as a by-product of microbial activity, and is commonly present in measurable amounts in the pore water of natural sediments. In some cases, ammonia concentrations can be high enough to be a potential influence on sediment test results, either through direct toxicity, or by inducing changes in organism behavior (for example, sediment avoidance; Whiteman et al. 1996 (41)). Depending on the nature of the assessment, toxicity caused by ammonia may be considered to be part of the aggregate effect of interest, while other assessments may be focused on more persistent chemicals (for example, metals, PAHs) and effects from ammonia may be viewed as an interference. The assessment and management of ammonia is discussed in greater detail within each individual test method (Annex A1 to Annex A5), including information on the sensitivity of each test organism to ammonia. Extending pre-test equilibration of test sediments may be useful in reducing ammonia concentrations in sediment through diffusion to and renewal of overlying water, provided it does not otherwise compromise study objectives. High concentrations of ammonia in field sediments may be indicative of nutrient enrichment; as the ecological effects of nutrient enrichment are far ranging, sediment toxicity testing may not be an appropriate tool for assessing nutrient enrichment as a stressor.

6.2.8 Low Dissolved Oxygen—Each of the test methods (Annex A1 to Annex A5) specifies a minimum dissolved oxygen concentration of 2.5 mg/L, a concentration that has been shown to be tolerated by all three of the test organisms described in this standard (Irving et al. 2004 (42), Mattson et al. 2009 (43)). Dissolved oxygen below 2.5 mg/L may reduce growth or survival. Because the food provided can add to oxygen demand, previous editions of this standard included direction that if dissolved oxygen approached 2.5 mg/L in any treatment, that feeding be reduced in all treatments, in an effort to maintain comparability among all treatments. In the current edition, a more nuanced approach to low dissolved oxygen is discussed, recognizing that having a small number of treatments compromised because of low dissolved oxygen may be preferable to altering the conditions (reduced food, aeration) in all treatments. Detailed discussion of this issue is contained within each individual test method in Annex A1 to Annex A5.

6.2.9 Influence of Indigenous Organisms—The potential effects of indigenous organisms, that is organisms present in the sediment at the time of collection, is difficult to study and not fully understood. Generally speaking, having sediments without other organisms present would allow a more standardized means of assessing the toxicity of sediment contaminants, but there is no widely accepted means of eliminating indigenous organisms from samples. Course sieving may be used to remove larger organisms, and extended storage under refrigeration may be effective in reducing the number of surviving indigenous organisms (10.3.2.3). Additional discussion is provided within each test method.

6.2.10 Influence of Light Quality—Some toxicants interact with certain wavelengths of light in a way that increases toxicity (for example, photo-induced toxicity of polycyclic aromatic hydrocarbons (PAHs); Davenport and Spacie 1991 (44), Ankley et al. 1994b (45)). Other toxicants may undergo photolysis if exposed to the appropriate wavelengths of light. Often such light-mediated effects are created by the higher-energy wavelengths present in sunlight, such as the ultraviolet range, which are generally not present to a similar extent in common fluorescent lighting as is typically used for laboratory sediment toxicity testing. Addressing this issue in the context of bedded sediments in the field is challenging, as the penetration of ultraviolet light can be greatly attenuated by the water column, and physical shading by sediment particles may further reduce exposure of benthic organisms and sediment associated. Where these processes may be important, study specific considerations will be required.

7. Water, Formulated Sediment, and Reagents

7.1 Water:

7.1.1 Performance-based Requirements:

7.1.1.1 Supporting Organism Performance—The primary requirement for water used to test and culture organisms is that it support satisfactory survival, growth, or reproduction of the test organisms. Test organisms should survive, grow, and reproduce at rates at or above those described in control performance criteria for long-term testing (for example, Table A2.1 and Table A4.1), and should not show signs of disease or apparent stress (for example, discoloration, unusual behavior).

Control tests used to demonstrate the adequacy of the water used to conduct a test should not be conducted using a natural sediment or other substrate that would contribute additional ions to the water; instead, a more inert substrate, such as rinsed sand, should be used to ensure that the test water alone is adequate to support organism health. Note that a demonstration of adequacy in long-term testing is required even if only 10-d tests are being conducted; this is important because some shortcomings of overlying waters only become apparent during longer exposures, even though they undoubtedly cause some degree of shorter-term impact on organism health. If problems are observed in the culturing or testing of organisms, it is desirable to evaluate the characteristics of the water. See Guide E729 for recommendations on chemical analysis of the water supply.

7.1.1.2 Bromide and Chloride Requirements for *Hyalella azteca*—Research has shown that the strain of *Hyalella azteca* most commonly cultured for laboratory testing has more specific requirements for bromide and chloride than does *Chironomus dilutus* (Borgmann 1996 (46); Soucek et al. 2015 (47), Ivey et al. 2016 (37), Ivey and Ingersoll 2016 (48)). Special attention to these constituents is required when selecting culture and test water for *H. azteca*.

(1) **Bromide**—Borgmann (1996) (46) reported the essentiality of bromide to *H. azteca*, and proposed a water containing 0.8 mg/L Br, a concentration well above that in most natural waters. Ivey and Ingersoll (2016) (48) conducting additional studies to refine the range of Br required for *H. azteca*, and concluded that a minimum of about 0.02 mg Br/L was required for culturing and testing of *H. azteca*, with poor survival, growth and reproduction of *H. azteca* observed in reconstituted waters having less than 0.02 mg/L Br. The adequacy of 0.02 mg Br/L was further supported in inter-laboratory studies (Ivey et al. 2016) (37), although a concentration of 0.04 mg Br/L is recommended to assure adequate bromide is present (Table A1.1 and Table A2.1). The requirement of *H. azteca* for Br appears to exist for multiple strains of this amphipod, as different strains were used by Borgmann (1996) (46) and Ivey et al. (2016) (37).

(2) **Chloride**—In addition to Br, the “US Lab” strain of *H. azteca* (Major et al. 2013) (49) has an additional requirement for Cl concentrations higher than is present in many natural or reconstituted waters. Soucek et al. (2015) (47) showed that performance of *H. azteca* declines when chloride concentrations are below about 15 mg/L (Soucek et al. 2015) (47). This is consistent with previous reports of difficulty culturing or testing *H. azteca* in reconstituted waters with lower Cl, such as that derived from the original formulation proposed by Marking and Dawson (1973) (50) and recommended in some toxicity testing methods for other species (USEPA 2002d) (51). In addition to reduced growth and reproduction, the “US Lab” strain of *H. azteca* also showed greater sensitivity to the acute toxic effects of sodium sulfate and sodium nitrate at lower Cl, in a concentration-depend manner similar to that observed for control growth and reproduction. In contrast, the growth of a genetically distinct strain of *H. azteca* obtained from an Environment Canada laboratory in Burlington, Ontario, Canada, was not influenced by chloride concentrations.

the acute toxicity of sodium sulfate or sodium nitrate. These findings suggest that the chloride-dependence of toxicity shown for the US Lab strain of *H. azteca* may be an unusual feature of that strain and perhaps not broadly representative of aquatic organisms as a whole.

(3) *Supplementing with NaBr and NaCl*—If a laboratory's control water does not meet or exceed both 0.04 mg Br/L and 15 mg Cl/L, the waters can be supplemented with NaBr and NaCl to reach these minima for any toxicity tests conducted with *H. azteca* (Annex A1 and Annex A3). A minimum of 0.04 mg Br/L is recommended because the studies of Ivey and Ingersoll (2016) did not determine where between 0.01 and 0.02 mg Br/L performance decreased, and spiking to a target Br concentration of 0.04 mg/L avoids being too close to the minimum requirement. Although not required, waters so adjusted appear completely compatible for testing with *C. dilutus* (Soucek et al. 2015 (47), USEPA Duluth laboratory and USGS Columbia laboratory, unpublished data).

7.1.2 Sources and Treatment:

7.1.2.1 *Natural Waters*—A natural water is considered to be of uniform quality if monthly ranges of the hardness, alkalinity, and specific conductance are less than about 10 % of their respective averages and if the monthly range of pH is less than about 0.5 units. Uncontaminated wells or springs are likely sources of suitable waters that are stable in composition. Surface waters may be used, but the intake should be positioned to: (1) minimize fluctuations in quality and contamination, (2) avoid depths with low oxygen content, and (3) ensure low concentrations of sulfide and iron. Surface waters should have low concentrations of dissolved organic carbon (for example, less than 5 mg/L), should be filtered to remove particulates.

7.1.2.2 *Municipal Tap Water*—Municipal water supplies may be temporally variable and may contain unacceptably high concentrations of materials such as copper, lead, zinc, fluoride, chlorine, or chloramines. Chlorinated water should not be used for culturing or testing because residual chlorine and chlorine-produced oxidants are toxic to many aquatic organisms. Use of tap water is discouraged unless it is rigorously dechlorinated (USEPA 2002d) (51).

7.1.2.3 *Reconstituted Water*—Reconstituted waters may be prepared by adding reagent grade salts to de-ionized water. Details and recipes are provided in 7.1.3.

7.1.2.4 *Adjusting or Matching Water Chemistry*—If desired, natural (or treated tap) waters may have their chemistry adjusted by diluting natural water with deionized water (for example, Kemble et al. 1994 (52), Besser et al. 2011 (53)) or by adding salts to relatively dilute natural waters (for example, Kunz et al. 2013 (54)), or a combination of both. For site-specific investigations, it may be desirable to have the water quality characteristics of the overlying water similar to the site water, provided those characteristics do not adversely affect the test organisms. Water chemistry is known to affect the water-column toxicity of certain contaminants common to sediments, such as many cationic metals, though the influence of overlying water chemistry on the toxicity of sediment-associated metals is less understood.

7.1.2.5 *Dissolved Gases*—Water might need aeration using air stones, surface aerators, or column aerators. Adequate aeration will stabilize pH, bring concentrations of dissolved oxygen and other gases into equilibrium with air, and minimize oxygen demand and concentrations of volatiles. Waters from sources that are substantially colder than test temperature, or that are collected at depth from surface waters (for example, >5 m), may contain dissolved gases (for example, oxygen, nitrogen) at concentrations substantially above those in equilibrium with ambient air and pressure. Because some gases can be slow to come to equilibrium, mechanical degassing equipment may be necessary to bring dissolved gases into an appropriate range.

7.1.3 Reconstituted Waters:

7.1.3.1 *Background*—Reconstituted waters offer the ability for laboratories with different water sources to conduct testing using the same water chemistry, and also provides a means to conduct tests in laboratories that do not have access to an appropriate source of natural water. Early in the development of sediment test methods, particularly for *H. azteca*, control performance problems were noted when tests were conducted with some common reconstituted waters, notably those derived from the formulation of Marking and Dawson (1973) (50), which has been incorporated into several standard toxicity test methods (for example, USEPA 2002d (51), Guide E729). As discussed in 7.1.1.2, insufficient concentrations of Br (<0.02 mg/L) or Cl (<15 mg/L) were likely a cause of many such problems with *H. azteca*. For tests conducted with the US Laboratory strain of *H. azteca* (Major et al. 2013) (49), reconstituted waters should contain at least 0.04 mg Br/L and 15 mg Cl/L. As emphasized in 7.1.1.1, reconstituted waters should only be used for sediment toxicity testing after they have been shown to support adequate survival, growth, and reproduction, such as in control tests meeting the control performance criteria for long-term tests (Table A2.1 and Table A4.1).

7.1.3.2 *Base Water for Preparing Reconstituted Waters*—Reconstituted water is generally prepared by adding specified amounts of reagent-grade chemicals to deionized water. Deionized water should be obtained from a system capable of producing at least 1 mega-ohm water. If large quantities of high quality deionized water are needed, it may be advisable to supply the laboratory grade water deionizer with water that has been pre-treated to reduce ion content, such as by a lower-specification mixed-bed ion exchange treatment or a reverse osmosis system, to extend the life of the laboratory grade system. Water prepared by reverse osmosis, or by distillation, may also be used provided it is shown to produce water of similar quality, and to be comparably free of contaminants. Note that many older water distillation systems contained components that would allow elevated concentrations of metals or other contaminants to be present in the product water.

7.1.3.3 *General Procedures for Preparing Reconstituted Waters*—Published recipes for reconstituted waters may have accompanying procedural instructions, though there is no evidence that the way in which the component salts are dissolved into de-ionized water is important. Some recipes dissolve salts individually into separate aliquots of water, while others combine them in different sub-

Calcium salts are sometimes dissolved in water separately from other salts (particularly bicarbonate (HCO_3)). All formulas recommended aeration for at least 24 h following salt addition to ensure complete dissolution and stabilization of pH. Salts should be sprinkled in gradually with mixing to avoid formation of a fused mass at the bottom of the container. In addition, dissolution of the salts should be visually verified, and further confirmed through characterization of the resulting water (for example, conductivity, hardness, alkalinity; 7.1.3.4).

7.1.3.4 Characterization of Reconstituted Waters—Conductivity, pH, hardness, dissolved oxygen, and alkalinity should be measured on each batch of reconstituted water. The reconstituted water should be aerated before use to stabilize pH and ensure dissolved oxygen is within acceptable ranges. USEPA (1991a) (55) recommends using a batch of reconstituted water for up to 2 weeks.

7.1.3.5 Examples of Reconstituted Fresh Waters—Information on several reconstituted waters that have been used successfully for sediment toxicity testing in at least some laboratories are provided below. No single water is uniquely recommended, and their order of appearance does not imply priority or preference. Aspects of different formulations that may affect water selection are noted. As per 7.1.1.1, any water selected needs to be shown to support organism health via control testing. It should be noted that most developmental work supporting the use of reconstituted waters in freshwater sediment toxicity test methods has been done in waters with a hardness at or near 100 mg/L as CaCO_3 , such as in the waters described in the subsections that follow; laboratories using waters with substantially different hardness (or otherwise different composition) should take extra care to ensure that any conclusions or recommendations developed based on waters in the 100 mg/L hardness range are applicable to other water chemistries.

(1) *Borgmann (1996) (46)*—Several laboratories have successfully used the formulation from Borgmann (1996) (46) to conduct water or sediment toxicity tests with *H. azteca* or with other test organisms. The Borgmann (1996) (46) water (“SAM-5S” shown in Table 5 of the original publication) includes the addition of 0.8 mg/L bromide; that concentration of bromide is much higher than that found in most natural waters and may be undesirable for that reason. However, laboratory testing has indicated that bromide is not harmful to *H. azteca* at concentrations as high as 20 mg/L (Ivey and Ingersoll 2016) (48). Later experiments by Ivey and Ingersoll (2016) (48) indicated that the Borgmann (1996) (46) water supported good survival, growth, and reproduction when the Br concentration was reduced to as low as 0.02 mg/L. For this reason, an alternative Br concentration of 0.04 mg/L is included here, which is closer to Br concentrations typical of North American surface waters, while still above the apparent minimum requirement (Table 2).

(2) *Smith et al., (1997) (56) with NaBr Amendment*—Smith et al. (1997) (56) reported on difficulties experienced with reconstituted water formulated after Marking and Dawson (1973) (50) and proposed a reformulation that both increased the Cl concentration of the water, and increased the Ca:Mg ratio, which is unusually low in the original Marking and Dawson (1973) (50) formula. The resulting formula is not

TABLE 2 Formula for Reconstituted Water after Borgmann (1996) (46) and Ivey and Ingersoll (2016) (48)

Salt	Concentration		Notes
	mM (millimolar)	mg/L	
$\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$	1	141	Other hydration states may be used with appropriate adjustment of concentration.
NaHCO_3	1	84.0	Other hydration states may be used with appropriate adjustment of concentration.
$\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$	0.25	61.6	
KCl	0.05	3.73	Top concentration is for 0.04 mg Br/L (Ivey and Ingersoll 2016) (48); bottom concentration is original Borgmann (1996) (46) formulation with 0.8 mg Br/L.
NaBr	0.0005 (or 0.01)	0.0519 (or 1.04)	

markedly different in composition from that proposed by Borgmann (1996) (46), with the exception that the original Smith et al. (1997) (56) formulation did not include any added Br, and some laboratories reported limited success with *H. azteca* in long-term testing using that original formulation. Ivey et al. (2016) (37) suggested that the absence of Br from the original formulation was the likely cause of those problems, and demonstrated good performance of *H. azteca* if Smith et al. (1997) (56) was supplemented to contain at least 0.02 mg Br/L. The formulation below contains 0.04 mg Br/L (Table 3). Note that Smith et al. (1997) (56) recommended dissolving the Ca salts in a separate aliquot of de-ionized water from the other salts, then combining the 2 solutions.

(3) *USEPA Duluth Soluble Salt Formula*—Because some existing reconstituted water formulas have ionic composition dissimilar from many natural waters, staff at the USEPA-Duluth laboratory developed a reconstituted water formulas intended to better match the characteristics of natural waters, based on surface water survey data collected by the U.S. Geological Survey (Alexander et al. 1996) (57). Two versions were developed, one that uses only easily dissolved salts (this section and Table 4) and another that uses CaCO_3 , which is more difficult to dissolve (7.1.3.5(4) and Table 5). While this

TABLE 3 Formula for Reconstituted Water after Smith et al. (1997) (56); with NaBr per Ivey and Ingersoll (2016) (48)

Salt	Concentration		Notes
	mM (millimolar)	mg/L	
$\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$	0.45	66.1	Other hydration states may be used with appropriate adjustment of concentration.
CaSO_4	0.37	50.0	Other hydration states may be used with appropriate adjustment of concentration.
NaHCO_3	1.14	96.0	
$\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$	0.25	61.6	
KCl	0.054	4.00	For 0.04 mg Br/L (Ivey and Ingersoll 2016) (48)
NaBr	0.0005	0.0519	



TABLE 4 Formula for Soluble Salt Reconstituted Water Developed by USEPA-Duluth (100 mg/L hardness)

Salt	Concentration		Notes
	mM (millimolar)	mg/L	
CaCl ₂ ·2H ₂ O	0.363	53.4	Other hydration states may be used with appropriate adjustment of concentration.
CaSO ₄ ·2H ₂ O	0.338	58.2	
NaHCO ₃	1.14	129	Other hydration states may be used with appropriate adjustment of concentration.
MgSO ₄ ·7H ₂ O	0.298	73.5	
KHCO ₃	0.0509	5.10	For 0.04 mg Br/L when testing <i>H. azteca</i> (Ivey and Ingersoll 2016) (48)
NaBr	0.0005	0.0519	

TABLE 5 Formula for Reconstituted Water with CO₂-aided Dissolution Developed by USEPA-Duluth (100 mg/L hardness)

Salt	Concentration		Notes
	mM (millimolar)	mg/L	
CaCO ₃	0.701	70.2	Other hydration states may be used with appropriate adjustment of concentration.
MgCl ₂ ·6H ₂ O	0.0655	13.2	
MgSO ₄ ·7H ₂ O	0.233	57.4	Other hydration states may be used with appropriate adjustment of concentration.
NaHCO ₃	0.342	28.7	
NaCl	.0755	4.41	If using water for <i>H. azteca</i> testing, also add the following:
KCl	.0509	3.80	
NaCl	0.165	9.67	For 0.04 mg Br/L (Ivey and Ingersoll 2016) (48))
NaBr	0.0005	0.0519	

soluble salt formula matches well with U.S. natural waters in several respects (for example, ratios of Ca:Mg, Cl:SO₄, hardness:alkalinity), the overall concentrations of Na+K are higher (relative to hardness) than in most surface waters. As for most reconstituted waters, this stems from adding alkalinity primarily as NaHCO₃, rather than from carbonates of Ca and Mg, as in most natural waters. CaCO₃ and MgCO₃, are not typically used in reconstituted waters because they dissolve extremely slowly if additional steps are not taken to dissolve them (for example, CO₂ enrichment; see 7.1.3.5(4)). Another consequence of adding alkalinity as NaHCO₃ (or KHCO₃) instead of CaCO₃ and MgCO₃ is that Ca and Mg must then be added as Cl or SO₄ salts, resulting in higher than typical concentrations of those anions. The formula developed by USEPA-Duluth reduces these problems by adding K as KHCO₃ instead of KCl (thus reducing NaHCO₃ addition), and by maintaining a typical Cl:SO₄ ratio even though Cl + SO₄ is high. When made to 100 mg/L hardness as shown below, this formula contains 25.7 mg Cl/L, thus exceeding the 15 mg/L minimum, and is supplemented with NaBr to achieve 0.04 mg

Br/L. This water has been used successfully in long term testing with the US Laboratory strain of *H. azteca* (Soucek et al. 2013 (58), 2016 (59), Ivey et al. 2016 (37)).

(4) *USEPA Duluth CO₂-aided Reconstituted Water*—As discussed in 7.1.3.5(3), closely matching the chemistry of typical natural waters generally requires that most alkalinity be added as CaCO₃ or MgCO₃. While adding alkalinity in this way allows creation of waters that have chemistry that better aligns with most natural waters, it is more difficult to prepare because these salts, particularly CaCO₃, dissolve very slowly at neutral pH. The dissolution can be speeded by temporarily reducing the pH by bubbling CO₂ through the solution. After the CaCO₃ has dissolved, the water can be reaerated with ambient air to bring the pCO₂ and pH back into normal ranges. Also, the original formulation of this water has a Cl concentration of about 9.1 mg/L less than the 15 mg Cl/L recommended for *H. azteca*. To use this formulation with *H. azteca*, additional NaCl can be added to raise the Cl concentration to 15 mg/L, as noted in the detailed directions (Table 5).

Procedure for preparing USEPA Duluth CO₂-aided reconstituted water:

(a) Add salts at the concentrations indicated in Table 5 to the appropriate volume of deionized water. If water is intended for use with *H. azteca*, add the additional amounts of NaCl and NaBr noted at the bottom of Table 5.

(b) In a hood, or other ventilated workspace that will ensure workers are not exposed to elevated CO₂ concentrations, bubble 100 % CO₂ gas through the solution using an airstone or other diffuser to disperse the gas. Vessels that have a taller geometry (higher than wide) will speed dissolution of the CO₂ and, as a result, the salts. Stir the solution as needed to aid dissolution of the salts. Time for dissolution will depend on the chamber geometry, gas flow, and bubble size, but dissolution should not usually take more than about 30 minutes.

(c) Bubbling with 100 % CO₂ will suppress pH considerably (for example, pH 5.0). After salts are dissolved, the solution is aerated with ambient air until appropriate pH (circa 6.8 to 8.3 is restored). Aeration may continue overnight if desired.

7.1.3.6 *Synthetic Seawater*—*Hyaella azteca* have been used to evaluate the toxicity of estuarine sediments using overlying waters up to 15 ‰ salinity in 10-d exposures (Nebeker and Miller 1988 (60), Roach et al. 1992 (61), Winger et al. 1993 (62), Ingersoll et al. 1996 (63)) and in 42-d exposures (Chris Ingersoll, USGS, Columbia, MO, unpublished data). It is important that high salinity testing be conducted using an ion mixture resembling seawater, rather than just NaCl, as *H. azteca* is considerably less tolerant of NaCl alone as compared to the broader mix of ions present in seawater; elevated K and Ca in seawater have been identified as important components of the increased tolerance of seawater (USEPA-Duluth, unpublished data). Introduction of seawater salts to *H. azteca* cultures should also be gradual, as sudden transfer to elevated salinity can induce toxicity (Soucek et al. 2013 (58), USEPA-Duluth, unpublished data). Ideally, the cultures of *H. azteca* should be slowly acclimated to elevated salinity.

before obtaining test organisms for conducting exposures in elevated salinity waters (7.1.3.6(2)).

(1) *Seawater Salts*—Reconstituted salt water can be prepared by adding commercial salt mixtures, such as FORTY FATHOMS®, HW MARINEMIX®, INSTANT OCEAN®, or equivalent to deionized water. A synthetic seawater formulation called GP2 is can also be prepared using reagent grade chemicals that can be diluted with deionized water to the desired salinity (USEPA 1994d) (64).

(2) *Culture at Elevated Salinity*—Ingersoll et al. (1992) (65) describe procedures for culturing *H. azteca* at salinities up to 15 ‰. Reconstituted salt water was prepared by adding INSTANT OCEAN® salts to a 25:75 (v/v) mixture of freshwater (hardness 283 mg/L as CaCO₃) and deionized water that was held at least 2 weeks before use. Synthetic seawater was conditioned by adding 6.2 mL of Fritzyme® #9 nitrifying bacteria (*Nitromonas* sp. and *Nitrobacter* sp.; Fritz Chemical Company, Dallas, TX) to each liter of water. The cultures were maintained by using renewal procedures; 25 % of the culture water was replaced weekly.

7.2 Formulated Sediment:

7.2.1 *Background and Purpose*—Formulated sediments are mixtures of materials that are intended to mimic the physical components of natural sediments (for example, a mixture of mineral particle sizes and containing a source of organic carbon). A primary use of formulated sediment could be as a control sediment. Formulated sediments allow for standardization of sediment testing or provide a basis for conducting sediment research. Formulated sediment provides a basis by which any testing program can assess the acceptability of their procedures and facilities. In addition, formulated sediment provides a consistent measure evaluating criteria required to determine test acceptability. The use of formulated sediment eliminates interferences caused by the presence of indigenous organisms. For toxicity tests with sediments spiked with specific chemicals, the use of a formulated sediment eliminates or controls the variation in sediment physico-chemical characteristics and provides a consistent method for evaluating the fate of chemicals in sediment.

7.2.2 *Considerations in Formulation*—A variety of recipes for creating formulated sediments were described in USEPA (2000 (8), 2019 (7)). A constraint on designing formulated sediments for broad use includes identifying materials that are broadly available, are representative of natural materials in field-collected sediments. For example, most all of the sources of silt, clay, and sand used to prepare formulated sediments described by Kemble et al. (1999) (66) are no longer commercially available. Moreover, the alpha cellulose used as a source of total organic carbon in the formulated sediment developed by Kemble et al. (1999) (66) has since been found to lack the binding characteristics and nutritional quality of naturally occurring sources of total organic carbon in natural sediments. Consistency of characteristics of the component materials is also important, meaning: (1) consistency of materials from batch to batch, (2) contaminant concentrations below concentrations of concern, and (3) availability to all individuals and facilities.

7.2.3 *Performance-based Acceptability*—Regardless of formulation, an essential characteristic of formulated sediments is that they be shown to support normal survival, growth, or reproduction for benthic invertebrates of interest. In the context of the toxicity methods described here, this generally means demonstrating that a control comprised of the formulated sediment can meet the control performance criteria for long-term sediment testing outlined in Annex A3 and Annex A4. Note that even if formulated sediments will only be used for 10-d sediment toxicity tests with *H. azteca* or *C. dilutus* (Annex A1 and Annex A2), the formulated sediment should still be shown to support adequate performance in longer term tests (Annex A3 and Annex A4) to ensure that the formulated sediment is fully capable of supporting organism health.

7.3 Reagents:

7.3.1 *Specifications*—Chemicals used as a part of sediment toxicity testing should be at least reagent grade, unless a test using a formulated commercial product, technical-grade, or use-grade material is specifically needed. Reagent containers should be dated when received from the supplier, and the shelf life of the reagent should not be exceeded. Working solutions should be dated when prepared and the recommended shelf life should not be exceeded.

8. Health, Safety, Waste Management, and Biosecurity

8.1 General Precautions:

8.1.1 *Scope*—Procedures described in this standard may involve hazardous materials, operations, and equipment, but it does not purport to address all of the safety problems associated with these activities. It is the responsibility of the user to establish appropriate safety and health practices, and determine the applicability of regulatory limitations before use. While some safety considerations are included, it is beyond the scope of this standard to encompass all safety requirements necessary to conduct sediment tests. Appropriate safety measures must be developed by each laboratory, and modified as necessary for studies posing unique hazards; the guidance in this section is only general.

8.1.2 *Components*—Development and maintenance of an effective health, safety, waste management, and biosecurity program in the laboratory requires an ongoing commitment by laboratory management and includes: (1) the appointment of a laboratory health and safety officer with the responsibility and authority to develop and maintain a safety program, (2) the preparation of a written health and safety plan, which is provided to each laboratory staff member, (3) an ongoing training program on laboratory safety, and (4) regular safety inspections.

8.1.3 *Diversity of Potential Risks*—Collection and use of sediment may involve substantial risks to personal safety and health. Contaminants in field-collected sediment may include carcinogens, mutagens, and other potentially toxic compounds. Inasmuch as sediment toxicity testing is often begun before chemical analyses can be completed, worker contact with sediment may need to be minimized by: (1) using gloves, laboratory coats, safety glasses, face shields, and respirators as appropriate, (2) manipulating sediment under a ventilated hood or in an enclosed glove box, and (3) enclosing and labeling

the exposure system. Personnel collecting sediment samples and conducting tests should take all safety precautions necessary for the prevention of bodily injury and illness that might result from contact with, or ingestion or inhalation of infectious agents, or corrosive or toxic substances, or asphyxiation because of lack of oxygen or presence of noxious gases.

8.1.4 *Preparation*—Before beginning sample collection and laboratory work, personnel should determine that all required safety equipment and materials have been obtained and are in good condition.

8.2 Safety Equipment:

8.2.1 *Personal Safety Gear*—Personnel should use appropriate safety equipment, such as laboratory coats, gloves, safety glasses, face shields, rubber aprons, respirators, and safety shoes.

8.2.2 *Laboratory Safety Equipment*—Each laboratory should be provided with safety equipment such as first aid kits, fire extinguishers, fire blankets, emergency showers, and eye-wash stations. All laboratories should be equipped with a telephone to enable personnel to summon help in case of emergency.

8.3 General Laboratory and Field Operations:

8.3.1 *Training and Education*—Laboratory personnel should be trained in proper practices for handling and using chemicals that are encountered during procedures described in this standard. Routinely encountered chemicals include acids, organic solvents, and standard materials for reference-toxicity tests. Special handling and precautionary guidance in Safety Data Sheets (formerly known as Material Safety Data Sheets) should be reviewed and followed for reagents and other chemicals purchased from supply houses.

8.3.2 *Work Practices*—Work with some sediment may require compliance with regulations pertaining to the handling of hazardous materials. Personnel collecting samples and performing tests should not work alone. It is advisable to wash exposed parts of the body with antibacterial soap and water immediately after collecting or manipulating sediment samples.

8.3.3 *Chemical Safety*—Strong acids and volatile organic solvents should be used in a fume hood or under an exhaust canopy over the work area. To prepare dilute acid solutions, the concentrated acid should be added to water, not vice versa. Opening a bottle of concentrated acid and adding concentrated acid to water should be performed only under a fume hood. An acidic solution should not be mixed with a hypochlorite solution because hazardous vapors might be produced.

8.3.4 *Electrical Safety*—Use of ground-fault systems and leak detectors is strongly recommended to help prevent electrical shocks. Electrical equipment or extension cords not bearing the approval of Underwriter Laboratories should not be used. Ground-fault interrupters should be installed in all “wet” laboratories where electrical equipment is used.

8.3.5 *Labeling*—All containers should be adequately labeled to identify their contents.

8.4 Disease Prevention:

8.4.1 *Vaccination and Hygiene*—Personnel handling samples that are known or suspected to contain human wastes

should be given the opportunity to be immunized against diseases such as hepatitis B, tetanus, typhoid fever, polio, and others as appropriate. Thorough washing of exposed skin with antibacterial soap should always follow handling these samples.

8.5 Pollution Prevention, Waste Management, and Sample Disposal:

8.5.1 *Compliance*—It is the laboratory’s responsibility to comply with the federal, state, and local regulations governing the waste management, particularly hazardous waste identification rules and land disposal restrictions, and to protect the air, water, and land by minimizing and controlling all releases from fume hoods and bench operations. In addition, compliance is required with any sewage discharge permits and regulations.

8.5.2 *Handling of Hazardous Materials*—Guidelines for the handling and disposal of hazardous materials should be followed. As a handler of hazardous materials, it is a laboratory’s responsibility to know and comply with the applicable federal and state regulations. Where possible, substitution of non-hazardous chemicals and reagents should be encouraged.

8.5.3 *Hazardous Waste*—Use of the methods described in this standard might result in wastes that are legally classified as “hazardous waste”, and thereby be subject to additional federal and state regulations.

8.6 Biosecurity and Non-native Species:

8.6.1 *Concerns for Non-native Species*—The introduction of non-native aquatic organisms from sources such as culturing facilities or toxicity testing facilities into a watershed is an environmental concern as invasive species can alter existing habitats (Zhu et al. 2006) (67), out-compete or replace native species (Meffe 1985 (68), Fausch 1988 (69), Walser et al. 2000 (70)), introduce diseases (Hoffman and Schubert 1984) (71), or change the community composition (Lemly 1985) (72).

8.6.2 *Sediment Toxicity Testing and Biosecurity*—Biosecurity represents precautions that can be taken to reduce the likelihood of releasing non-native organisms or introduction of diseases associated with toxicity testing. Test organisms used to conduct a laboratory toxicity test may not be native to the watershed where testing is conducted. Field-collected sediments or effluent generated from laboratory testing have the potential to contain organisms that are not native to the watershed where laboratory testing is conducted.

8.6.3 *Minimizing Potential for Release from the Laboratory*—The potential for release of non-native organisms from a culturing or testing laboratory can be reduced by using physical or chemical treatment of water and sediments from sediment toxicity testing. For example, effluent water generated from laboratory testing can be passed through 100-µm mesh opening polypropylene screen to capture macroinvertebrates inadvertently released from the sediment (for example, during the exposures and during the sieving of sediments at the end of the exposures; see Ingersoll et al. 2013 (73)). Maximizing surface area can reduce clogging of the mesh screens (for example, Fig. 1). Effluents generated from laboratory testing can also be treated with chlorine to kill non-native organisms or to help destroy disease organisms (for example, Fig. 2). Effluent generated from sieving sediment should be disposed

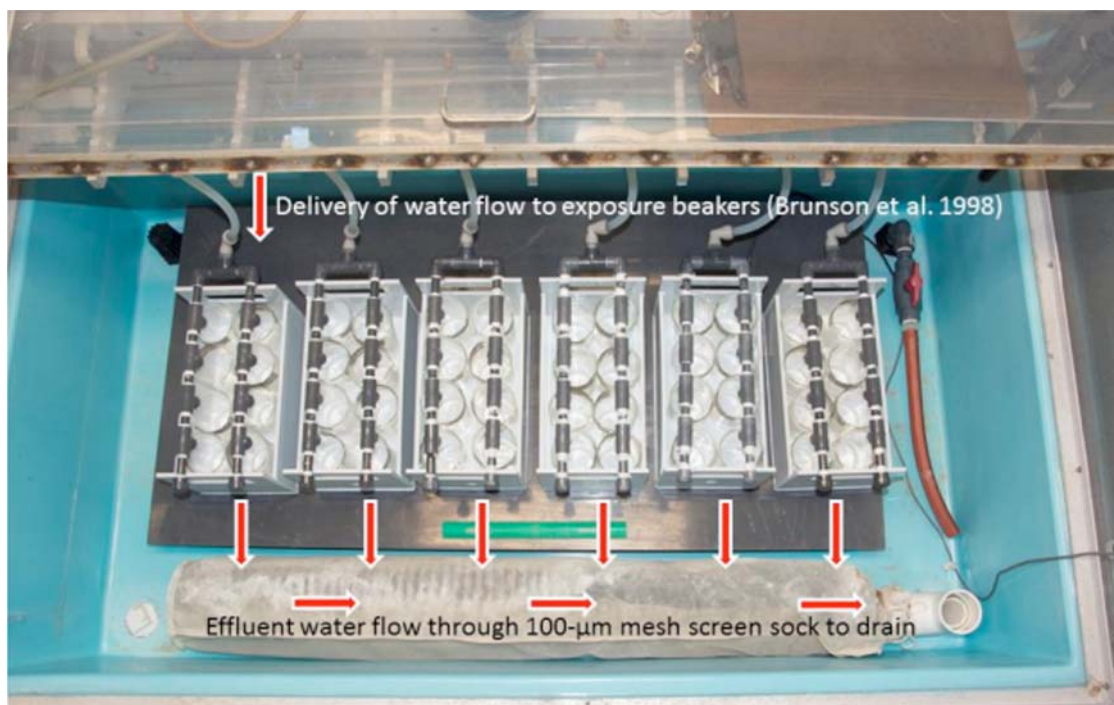


FIG. 1 An Example of a 100-µm Mesh Screen Sock Surrounding a Perforated Drain Tile Used to Collect Effluent Generated in a Sediment Toxicity Testing Water Bath

exposure can also be collected and passed through a 100-µm mesh screen sock before the effluent is released to a drainage system in a testing laboratory. While sediments used in toxicity testing will generally be disposed of by means that would limit the release of non-native aquatic species (for example, landfill), freezing sediments remaining at the end of a laboratory test may help destroy non-native organisms.

8.6.4 Non-native Species Concerns During Field Collection—Equipment and materials used for field collection of sediment samples could inadvertently transfer non-native species among waterbodies. Local resource management officials should be contacted before field sampling to identify any particular concerns.

9. Facilities, Equipment, and Supplies

9.1 General—This section contains information to inform development of adequate facilities, equipment and supplies for conducting the procedures described in this standard; it is not necessarily comprehensive and in some cases alternative items with comparable performance may be suitable. However, the adequacy of a laboratory and its material and equipment should be determined not simply through compliance with individual specifications, but also by the ability to support conduct of the described procedures in a manner that meets performance requirements. Before a toxicity test is conducted in a facility, it is strongly recommended that laboratories conduct a “control” test with each test species in which all test chambers contain control water and a control sediment (11.3.3). Survival, growth, or reproduction of the test organisms will demonstrate

whether facilities, water, diet, control sediment, and handling techniques are adequate to meet test acceptability criteria.

9.2 Facilities:

9.2.1 General Facilities for Culturing and Testing—The facility should include separate areas for culturing test organisms and for conducting water or sediment toxicity testing to reduce the possibility of contamination by test materials and other substances, especially volatile compounds. Holding, acclimation, and culture chambers should not be in an area where water or sediment tests are conducted, stock solutions or sediments are prepared, or equipment is cleaned. Test chambers may be placed in a temperature-controlled recirculating water bath, environmental chamber, or equivalent facility with temperature and lighting control. The water in a water bath should be continuously circulated to ensure mixing and provide uniform temperature around all test chambers. An enclosed and ventilated test system is desirable to limit exposure of laboratory personnel to volatile substances and to minimize transfer of such substances among test chambers.

9.2.2 Light—Light of the quality and luminance normally obtained in the laboratory is adequate (about 100 to 1000 lux using wide-spectrum fluorescent lights; for example, cool-white or daylight) has been used successfully to culture and test organisms. Lux is the unit selected for reporting luminance in this standard. Multiply units of lux by 0.093 to convert to units of footcandles. Multiply units of lux by 6.91×10^{-3} to convert to units of $\text{E/m}^2/\text{s}$ (assuming an average wavelength of 550 nm ($\text{mol}^{-2}\text{s}^{-1} = \text{W m} \times \lambda [\text{nm}] \times 8.36 \times 10^{-3}$); Guide E1733). Luminance should be measured at the surface of the water in



NOTE 1—Chlorine concentrations in the reservoir are maintained using chlorine pool tablets. For example, chlorine concentrations of >60 mg /L of total residual chlorine were maintained in a 35-L reservoir containing 3 chlorine tablets that received 16 volume additions/d of clean water from the diluter (for example, the 12-L diluter cycling every 30 min [average water addition of 400 mL/min]; Ingersoll et al. (2013) (73). Chlorine tablets should be replaced when less than 50 % of the tablets are remaining in the reservoir. A meshed screen sock can also be placed in the chlorine reservoir to increase residence time of organisms released into the effluent.

FIG. 2 Example of a Chlorine Reservoir Described in Ingersoll et al. (2013) (73) Used to Maintain a Target Concentration of Chlorine in the Effluent Generated from a Laboratory Toxicity Testing Diluter (a yard stick is provided for scale)

test chambers. A uniform photoperiod of 16L:8D (light:dark) can be achieved in the laboratory or in an environmental chamber using automatic timers.

9.2.3 Isolation from Disturbance—During phases of culturing, holding, and testing, test organisms should be limited from external disturbances such as rapidly changing light, temperature, or pedestrian traffic.

9.2.4 Ventilation—The test facility should be well ventilated and free of fumes. Laboratory ventilation systems should be checked to ensure that return air from chemistry laboratories or sample handling areas are not circulated to culture areas or testing areas, or that air from testing areas does not contaminate culture areas. Air pressure differentials between areas

should not result in a net flow of potentially contaminated air to sensitive areas through open or loose-fitting doors.

9.2.5 Air Supply for Aerating Water—Air used for aeration must be free of oil and fumes. Air blowers (high volume, low pressure) may be preferable to air pumps (high pressure, low volume), given that air blowers may be less expensive and easier to maintain. Oil-free air blowers or air pumps should be used. Filters to remove oil, water, and bacteria from the air supplied to culture or test organisms are desirable. High efficiency (for example, 99.99 %) coalescing filters are recommended (for example, BALSTON® Grade BX; Balston, Inc., Lexington, MA), and oil and other organic vapors can be

removed using activated carbon filters (for example, BAL-
STON® C-1 filter), or equivalent.

9.3 *Compatible Materials:*

9.3.1 *Acceptable and Prohibited Materials: Preparation—*Equipment and supplies that contact stock solutions, sediment, or overlying water should not contain substances that can be leached or dissolved in amounts that adversely affect the test organisms. In addition, equipment and supplies that contact sediment or water should be chosen to minimize sorption of test materials from water or sediment. Glass, type 316 stainless steel, nylon, high-density polyethylene, polypropylene, polycarbonate, and fluorocarbon plastics should be used whenever possible to minimize leaching, dissolution, and sorption of chemicals. Concrete and high-density plastic containers may be used for holding and culture chambers, and in the water-supply system. These materials should be washed in detergent, acid rinsed, and soaked in flowing water for at least a week before use. Cast-iron pipe should not be used in water-supply systems because colloidal iron will be added to the overlying water and strainers will be needed to remove rust particles. Copper, brass, lead, galvanized metal, and natural rubber must not contact overlying water or stock solutions before or during a test. Items made of neoprene rubber and other materials not mentioned above should not be used unless it has been shown that their use will not adversely affect survival, growth, or reproduction of the test organisms.

9.3.2 *Evaluation of Plastics—*New lots of plastic products should be tested for toxicity by exposing organisms to them under ordinary test conditions before general use.

9.4 *List of General Equipment:*

9.4.1 *Environmental Chamber—*Environmental chamber or equivalent facility with photoperiod and temperature control (for example, 20 °C to 25 °C).

9.4.2 *High Purity Water System—*Water purification system capable of producing 1 mega-ohm water or better (USEPA 1991a) (55).

9.4.3 *Analytical Balance—*Analytical balance capable of accurately weighing to 0.01 mg.

9.4.4 *Reference Weights—*Reference weights, Class S: for documenting the performance of the analytical balance(s). The balance(s) should be checked with reference weights that are at the upper and lower ends of the range of the weighings made when the balance is used. A balance should be checked at the beginning of each series of weighings; additional checks during, or at the end of a series may be desirable, particularly if large numbers of samples are being weighed. If desired, lower class reference weights (for example, Class F) may be used for routine monitoring of calibration, provided these weights are periodically compared to results with Class S weights. Blank weighing pans (not containing test organisms and processed consistently with all weigh pans) should be included with each batch of test organisms weighed.

9.4.5 *Volumetric Glassware—*Volumetric flasks and graduated cylinders—Borosilicate glass or nontoxic plastic lab ware, 10 to 1000 mL as needed for making test solutions.

9.4.6 *Pipets—*Volumetric pipets—Borosilicate glass, nontoxic plastic pipets (for example, 1 to 100 mL); bulbs and fillers as needed.

9.4.7 *Repeating Pipettor—*Adjustable-volume repeating pipette for providing food to test organisms (for example, repeated delivery of 1 mL from a 50-mL reservoir; not essential but very useful).

9.4.8 *Droppers or Similar Glass Tubing—*Droppers (glass or plastic) or glass tubing with fire polished edges, 4- to 6-mm ID—for transferring test organisms.

9.4.9 *Wash Bottles—*Wash bottles for rinsing small glassware, instrument electrodes and probes.

9.4.10 *Thermometer—*Glass or electronic thermometer—for periodic or continuously measuring water temperature (± 0.1 °C).

9.4.11 *NBS Thermometer—*National Bureau of Standards Certified thermometer (see USEPA Method 170.1; USEPA 1979b (74)), for verifying accuracy of routinely used thermometers.

9.4.12 *Meters for Wet Chemistry—*Dissolved oxygen (DO), pH/selective ion, and specific conductivity meters and probes for routine physical and chemical measurements are needed. Unless a test is being conducted to specifically measure the effect of DO or conductivity, a portable field-grade instrument is acceptable.

9.4.13 *Additional Details—*USEPA (2000) (8) provides a list of additional equipment and supplies for culturing or testing waters or sediment (for example, Table 6.1 in USEPA 2000 (8) and Appendix C in USEPA 2000 (8)).

9.5 *Exposure System:*

9.5.1 *Automated Water Delivery Systems—*Appendix A in USEPA (2000) (8) provides examples of water-renewal systems that have been used to conduct water or sediment toxicity testing. The system should be capable of delivering water to each replicate test chamber. A variety of toxicant dilution systems have been successfully modified for sediment testing. (Mount and Brungs (1967) (75), Ingersoll and Nelson 1990 (76), Maki 1977 (77), Benoit et al. 1993 (78), Zumwalt et al. 1994 (79), Brunson et al. 1998 (80), Brunson and Ingersoll 2006 (81), Besser et al. 2011 (53), Kunz et al. 2019 (82)). The water-delivery system should be capable of delivering a relatively consistent volume of water to replicate exposure chambers (for example, within about 10 % across replicates). The general operation of the system should be visually checked daily throughout the length of the test. If necessary, the water-delivery system should be adjusted during the test.

9.5.2 *Manual Water Renewal—*The overlying water can be replaced manually (for example, siphoning); however, manual systems take more time to maintain during a test. In addition, automated systems generally result in less suspension of sediment compared to manual renewal.

9.5.3 *Test Chambers—*Test chambers may be constructed in several ways and of various materials, depending on the experimental design and the contaminants of interest. Clear silicone adhesives, suitable for aquaria, sorb some organic compounds that might be difficult to remove. Therefore, as little adhesive as possible should be in contact with the test material. Extra beads of adhesive should be on the outside of the test chambers rather than on the inside. To leach potentially toxic compounds from the adhesive, all new test chambers constructed using silicone adhesives should be soaked in water for 48

hours in overlying water before use in a test. Test chambers for specific tests are described in [Annex A1](#) to [Annex A8](#).

9.5.3.1 Appendix A in USEPA (2000) (8) provides examples of test chambers. In addition, Brunson and Ingersoll (2006) (81) describe a refinement to the Brunson et al. (1998) (80) water-delivery system with an 8-way splitter and associated beaker box with integrated supports for the splitter ([Fig. 3](#)) to replace Zumwalt et al. (1994) (79) water-delivery system for conducting sediment toxicity tests. Compared to the Zumwalt et al. (1994) system, the refined 8-way splitter illustrated in [Fig. 3](#) has several advantages including: (1) ease of construction, (2) less space needed in exposure system (3) easier access to exposure chambers during exposures with less disturbance to test organisms, (4) increased durability, (5) easier to store in less space, (6) no wasted water, (7) a more consistent split of the water over the course of an exposure (for example, delivery to 48 beakers typically varies by <10 %), (8) much less maintenance of the system during an exposure, (9) easier to modify for different applications, and (10) the design of the system does not allow water exchange between replicate exposure chambers ([Guide E1241](#)). This 8-way splitter system has been used in numerous sediment and water exposures with various species including *H. azteca* and *C. dilutus*, and freshwater mussels (Besser et al. 2013 (83), 2015 (84); Ingersoll et al. 2015 (85); Ivey et al. 2016 (37); Kunz et al. 2017 (86); Wang et al. 2013 (87), 2017 (88)).

9.5.3.2 The 8-way splitter was designed using threaded 1/4-inch PVC pipe and fittings, silicone stoppers, and modified hypodermic needles to deliver about 125 mL/cycle to each of 8 replicate exposure chambers. The splitter is connected by flexible tubing to an elevated chamber or head box that receives about 1 L per splitter per cycle. The splitter sits on supports incorporated into the ends of the box containing the exposure beakers. When on the supports, each needle is suspended directly over and about 1 inch above 1 of the 8 beakers. The manifold can be easily lifted off the supports on the beaker box and set aside for full access to the beakers. The supports incorporated into the beaker box also act as handles to lift and carry the box by. Many procedures such as feeding, water sampling and general observations can be done without first removing the splitter.

9.5.3.3 The functional design of the splitter is based on the use of water head pressure to feed the splitter along with the uniform constriction in the flow path to each of the 8 beakers. The modified hypodermic needles used to constrict flow are cut to uniform length and cleaned of burrs to produce even flow to each beaker. In addition to enabling even flow through the splitter, the pressure of the water helps clear out debris and therefor helps keep the needles clean over time. In addition to requiring some water head pressure to function properly, the water needs to be delivered to the splitter as a pulse. If water slowly dribbles into the splitter, then the flow to each beaker

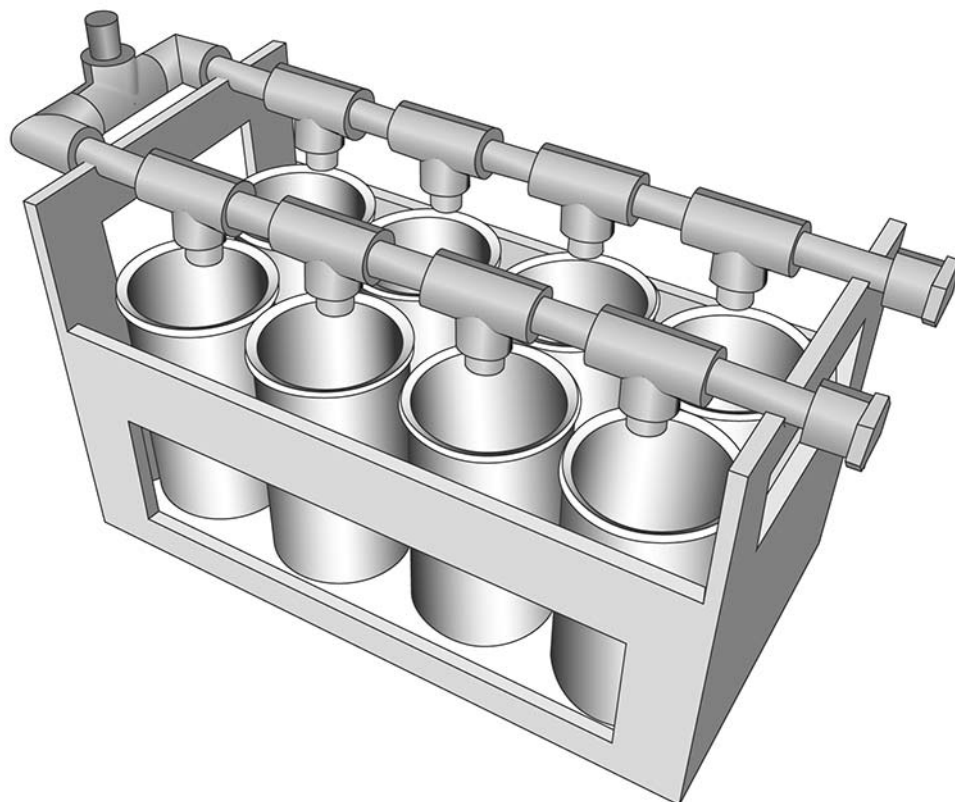


FIG. 3 Diagram of the 8-way Splitter and Associated Beaker Box with Integrated Supports for the Splitter Used to Deliver Overlying Water in Sediment Exposures (Brunson and Ingersoll 2006 (81), Kunz et al. 2019 (82))

will not be uniform and most of the flow will go to only a few of the beakers. The pulse of water can be provided by an automated water-delivery system modified from Mount and Brungs (1967) (75) proportional diluter (Ingersoll and Nelson 1990) (76), or manually by pouring a measured volume of water into the elevated chamber (or funnel) that the splitter is connected to. The stream of water produced by the splitter is strong enough to reach the bottom of a filled 300-ml tall form beaker but not so strong as to greatly disturb sediment during an exposure. The stream of water helps to mix the water column providing fresh water and food throughout the beaker. This provides improved water quality including elimination of much of the ammonia that can build up over time.

9.5.3.4 Construction of the splitter is relatively simple with all parts “off the shelf” except for the stoppers with needle inserts. Construction consists of wrapping Teflon thread tape onto the pipe fittings and threading them together making sure that the spacing between needles matches the spacing of the beakers below. The silicone stoppers (size 00) with needle inserts are made by cutting the beveled point off of a 16 gauge by 1.5-inch long hypodermic needle, inserting the needle through the center of the stopper, cutting off the hub-end, cleaning out the bore with a piece of stainless-steel wire, and rounding out each cut end with a stainless-steel dental hook. The needles can be cut with a Dremel tool using a reinforced cut-off wheel. Insertion through the stopper is made easier if a slight bevel is ground around the entire edge of the needle after the point is cut off. Table 6 provides a list of parts needed to construct an 8-way splitter.

9.5.3.5 The beaker boxes with integrated supports for the splitter are constructed of welded ¼-inch thick polypropylene sheeting. The design incorporates built in handles for moving the boxes with the 8 beakers contained inside. The sides of the boxes have panels cut out to allow for water circulation around the beaker while in a temperature-controlled water bath. The open sides also allow water to rapidly drain out of the beaker box when it is lifted out of the water bath. Although this design has been proven to function well in numerous studies, a different solution for suspending the splitters over the receiving beakers could be used. The main consideration in other designs is that the splitter has a support that holds it in place above the beakers and allows the splitter to be easily removed. When in use, the aim of each stream of water from the splitter should be checked to make sure it is pointed into the beaker and adjusted

if necessary. The aim should be checked each time a stopper/needle is removed and replaced for cleaning or inspection. Since the splitters are constructed with threaded fittings it is possible that with handling or during cleaning the position of the fittings can get out of alignment. When placing the needled stoppers in the splitter, it is important that they be screwed in until they hit the ledge at the end of the threads and can’t be screwed in deeper. If not done properly, the stopper may come out during a cycle allowing the great majority of water to go into the associated beaker.

9.6 *Cleaning*—All non-disposable sample containers, test chambers, and other equipment should be washed after use. For example:

- (1) Soak 15 min in tap water and scrub with detergent, or clean in an automatic dishwasher.
- (2) Rinse 3 times with tap water.
- (3) Carefully rinse once with fresh, dilute acid (for example, 10 % V:V or hydrochloric or nitric acid) to remove scale, metals, and bases.
- (4) Rinse 3 times with deionized water.
- (5) Rinse once with full-strength acetone to remove organic compounds (use a fume hood or canopy). Hexane might also be used as a solvent for removing nonionic organic compounds. However, acetone is preferable if only 1 organic solvent is used to clean equipment.
- (6) Rinse 3 times with deionized water and dry in a manner that will limit potential for recontamination.

10. Sample Collection, Storage, Characterization, and Spiking

10.1 *Collection*—The procedures for collecting, storing, manipulating, and characterizing sediments for sediment toxicity testing are intrinsically tied to the purpose of the study, and it is critical that procedures to be used are carefully considered in the context of study goals. Both Guide E1391 and USEPA (2001) (89) have produced detailed guidance on the collection, storage, and manipulation of sediments, and the reader is referred to those documents for detailed guidance on planning and conducting these activities. The guidance in this section represents only synoptic discussion of important issues, along with some more recent or specific information that is particularly applicable to sediment toxicity testing since the publication of USEPA (2000 (8), 2001 (89)).

10.2 *Collection of Sediments from the Field*:

10.2.1 *Safety and Disposal*—Field-collected sediments may contain multiple contaminants (or biological components) that can present hazards to both workers and the environment. Before the collection of field sediments, procedures for handling, storage and disposal of sediment should be established that are appropriate for the chemicals (or biological hazards) that may be present.

10.2.2 *Sampling Plan*—It is important that the approach to sediment collection be consistent with the study objectives (Guide E1391 and USEPA 2001 (89)). Sediments in the environment can show considerable spatial and temporal variability, and careful consideration should be given to how this variability will be addressed by sediment sampling in terms of the type of collection equipment used and the sampling

TABLE 6 Materials Needed for Constructing an 8-way Splitter (Brunson et al. 2006 (81), Kunz et al. 2019 (82))

Parts Description	Number Required for a Splitter
Tee	9
90° elbow	2
Close nipple (all thread)	2
38.1 mm (1.5 in.) nipple	6
50.8 mm (2 in.) nipple	4
Cap	2
Acetal instant tube fitting – swivel 90° elbow 6.35 mm (0.25 in.) pipe thread × 12.7 mm (0.5 in.; outside diameter) tubing	1
Silicone stoppers (size 00)	8
16 gauge × 38.1 mm (1.5 in.) hypodermic needles	8
Teflon pipe thread tape	1

of individual grabs from a single location. The volume of sediment required for all physical and chemical characterizations, as well as for toxicity testing need to be considered in light of sampling equipment and other logistical factors. In some cases, conditions at a particular field location may be incompatible with collection of an appropriate sample using the planned sampling equipment (for example, rock or cobble substrate interfering with core or dredge penetration); the sampling plan should include alternatives to be implemented in such cases.

10.2.3 Non-indigenous Organisms—Sediment sampling plans should include explicit consideration of the potential for the spread of non-indigenous organisms, either brought to the sampling site, or transferred from the sampling site via samples, sampling equipment, boats, personal gear, etc. See 8.6 for additional information on biosecurity.

10.2.4 Sampling Equipment—A wide variety of sediment sampling equipment exists, each having different attributes in terms of the sediment volume, sampling depth, degree of sediment disturbance, and effectiveness in different substrate types (Guide E1391 and USEPA 2001 (89)). The choice of sampling equipment will require balancing these characteristics against the requirements of the study. Many sediments have distinct stratigraphy and are therefore not homogenous with depth; if only certain sediment layers are to be assessed, sample collection equipment must be carefully chosen to preserve stratigraphy sufficiently for the study purposes. It is common to restrict the depth of sample collection to focus on the depths to which naturally occurring benthic organisms would inhabit. This depth depends on the types of invertebrates present, but it is common to focus on the upper 2 to 10 cm of sediment; again, this should be considered as part of the study planning process.

10.2.5 Replication and Compositing—It is common for the required sediment sample volume to exceed the capacity of the sediment collection equipment in a single grab, thus requiring compositing of multiple grabs into a single sample. Compositing of subsamples collected over a defined area may also be desirable if the study objectives are focused more on areal aerial average exposure than on fine scale heterogeneity. If both aspects are important, a combined approach can be used wherein individual grabs are subsampled for characterization of fine scale variability (typically for key chemical characteristics) while composites of these grabs are used for other measures (such as sediment toxicity testing).

10.2.6 Large Debris—It is common to encounter large debris (for example leaves, rocks, wood) during sediment sampling. Coarse screening (for example, 1 cm or larger) of sediments in the field may be desirable to eliminate this material and obtain a more uniform size distribution of sediment particles for conducting laboratory toxicity testing. Given the relatively uncontrolled conditions common to field sampling, and the general desire to minimize handling and disruption of the samples, it is typically desirable to limit sieving of sediment samples to relative coarse mesh sizes (for example, 1 cm or larger) through which the samples will pass with minimal effort, and without adding additional water. Removing larger debris by hand may also be sufficient if there

are only small amounts to be removed. Annex A1 to Annex A5 discuss sieving of sediment samples to smaller particle sizes for sediment toxicity testing, depending on study objectives; this is generally better pursued after initial sample collection to minimize the effort required in the field and allow the finer sieving to be conducted under more controlled conditions.

10.2.7 Protecting Sample Integrity—Some sediment contaminants, such as cationic metals, have chemical speciation tied closely to redox conditions, which can be disrupted during sediment collection and handling; accordingly, the degree of sample manipulation in the field should be minimized. Exposure to direct sunlight during collection should be minimized, especially if the sediment contains compounds that may degrade or transform in sunlight. Sediment samples should be cooled to 4 °C as soon as possible before shipment (Guide E1391). Ice can be used to cool samples in the field (for example, placing sample containers in a cooler containing ice). Refrigeration trucks or trailers can also be used to cool samples before transport (for example, Wang et al. 2013 (87)). Sediments to be used in laboratory toxicity tests should not be frozen because freezing will likely disrupt the physical and chemical characteristics of the sediment. Maximum/minimum thermometers or electronic data loggers can be used to monitor temperature during shipping and storage (for example, Hobo data loggers, Onset Computer Corporation, Bourne, MA USA; <http://www.onsetcomp.com/products/data-loggers-sensors/temperature>).

10.2.8 Indigenous Organisms—The potential influence of indigenous organisms is discussed as part of the toxicity test methods in Annex A1 to Annex A4. In general, it is not feasible to remove indigenous organism during field collection, except to the extent that they are removed by coarse sieving (10.2.6). Large organisms can be removed with forceps if practical. Field notes on the presence of indigenous organisms may be useful to inform later testing of the samples. Refrigerated storage of samples may reduce the survival of indigenous organisms.

10.3 Storage:

10.3.1 Storage Conditions—Unless there is a specific reason to do otherwise, sediments should be stored in the dark under refrigeration (at about 4 °C). Refrigeration slows the rate of microbial activity that may alter sediment characteristics, or increase concentrations of ammonia, sulfide, or other byproducts of microbial activity. Freezing of sediments should be avoided, as freezing can alter the texture of sediments and may also alter chemical partitioning (Guide E1391, Schuytema et al. 1988 (90)). Sediment may be stored in containers constructed of materials appropriate to the types of contaminants likely to be present (Section 9). Most samples of fine-grained sediments are comprised largely of anoxic sediment, and should be held with minimum headspace or a headspace purged with nitrogen to limit oxygen supply.

10.3.2 Holding Time:

10.3.2.1 General—Previous guidance on the allowable holding time before testing has varied, with recommended storage times ranging from 2 weeks (Guide E1391) to less than 2 months (USEPA-USACE 1998a) (91). Studies have been conducted to address issues related to sediment storage times

et al. 1994 (92), Becker et al. 1995 (93), Carr and Chapman 1995 (94), Moore et al. 1996 (95), Sarda and Burton 1995 (96), Sijm et al. 1997 (97), DeFoe and Ankley 1998 (98), Simpson and Batley 2003 (99)). The conclusions and recommendations offered by these studies depend primarily on the type or class of contaminant(s) present. Sediments that exhibit low-level to moderate toxicity can exhibit considerable temporal variability in toxicity, although the direction of change is often unpredictable (for example, Carr and Chapman 1995 (94), Moore et al. 1996 (95), DeFoe and Ankley 1998 (98)). Considered collectively, these studies suggest that a default recommendation to conduct sediment toxicity testing sometime between the time of collection and 2 months storage is reasonable.

10.3.2.2 Contaminant-specific Holding Time—While a 2-month holding time is reasonable as a default, alternative holding times, shorter or longer, may be appropriate for a particular study, depending on the contaminants involved and the study objectives. Sediments that contain comparatively stable compounds (for example, high molecular weight compounds such as polychlorinated biphenyls (PCBs), polycyclic aromatic hydrocarbons (PAHs), or oil) or which exhibit a moderate-to-high level of toxicity, typically do not vary appreciably in toxicity in relation to storage duration of months to even years (Moore et al. 1996 (95), DeFoe and Ankley 1998 (98), Ingersoll et al. 2014 (100)). As an extreme case, samples contaminated with PAHs have been held at US EPA Duluth laboratory for more than 10 years without notable change in toxicity (US EPA-Duluth, unpublished data). In contrast, other compounds may be more labile (for example, ammonia) or easily degraded, and may be less stable in storage. Sarda and Burton (1995) (96) found decreases in ammonia concentrations in pore water of sediments stored longer than 2 weeks warranting greater concern about limiting holding time to less than 2 months. Chemicals whose presence in sediment is the result of episodic events, such as pesticide runoff, may degrade relatively quickly in sediments. While metals do not degrade and therefore total metal concentrations should not change during storage, the bioavailability of metals is dependent on several sediment characteristics that may not be stable, such as acid-volatile sulfide (AVS).

10.3.2.3 Effects on Indigenous Organisms—Storage of sediment can reduce the viability of indigenous invertebrates present in field sediments. DeFoe and Ankley (1998) (98) observed high variability in survival during early testing periods (for example, <2 weeks) in sediments with low-level toxicity. They hypothesized that this variability partially reflected the presence of indigenous predators that remained alive when sediments were tested immediately or after short storage periods. Their data suggest, but do not prove, that it may be desirable to store sediment before testing (for example, >1 month) to reduce potential for interferences from indigenous organisms.

10.3.2.4 Monitoring Chemical Stability During Storage—Particularly if sediments are stored longer than generally recommended, it may be desirable to document any changes in sediment chemistry that might occur during storage. Depending on the nature of the sediment contaminant(s), parameters that might be monitored include contaminant concentration in

bulk sediment, concentration in pore water, or important ancillary measures such as AVS. Metal concentrations in pore water have been observed to change considerably during periods of storage (Simpson and Batley 2003) (99). Concentrations of chemicals of concern could be measured periodically in pore water during the storage period and at the start of the sediment test (Kemble et al. 1994 (52), Ingersoll et al. 2014 (100)).

10.4 Sediment Manipulation—Homogenization, Pore Water Isolation:

10.4.1 Homogenization—Sediment samples tend to settle during shipment or storage, typically resulting in a layering of more densely packed particles at the bottom and a watery layer at the top. This water layer should be considered as part of the sediment sample, and should not be discarded. Sediments should be homogenized before subsamples are taken for sediment chemical or physical analyses or before sediments are placed in exposure chambers. Homogenization of sediment can be accomplished using a variable-speed hand-held drill outfitted with a stainless-steel auger, propeller-type paint mixer, or similar device. Mixers of 5 to 8 cm in diameter are generally effective for containers about 19 L; smaller or larger mixers are available for larger or smaller containers. Mixing speed should be sufficient to produce a visually homogenous mixture (usually less than 5 minutes), but should not be so vigorous as to entrain air. Sediments can also be homogenized manually with a stainless steel spoon or paddle, but for larger volumes this can require considerable time and effort to achieve homogeneity.

10.4.2 Sediment Dilution—In some cases, diluting contaminated sediment with uncontaminated sediment may be desirable, such as to evaluate the relative potency of a sediment that kills most or all organisms in its undiluted form (for example, Ingersoll et al. 2015 (85)), or to reduce the potency of a sediment for use in sediment toxicity identification studies (USEPA 2007) (6). Predicting or interpreting the effect of sediment dilution is not as straightforward as for diluting water samples, as the effect of dilution will likely not be proportional to the degree of dilution based on volume or mass; this is because of the many sediment phases that may influence the bioavailability of sediment-associated contaminants. For example, the bioavailability of non-polar organic compounds is influenced primarily by the organic carbon fraction of the sediment (Di Toro et al. 1991) (2). As such, all other sediment characteristics being equal, if the contaminated sediment has 4 % organic carbon and the diluent sediment has 1 % organic carbon, a 50:50 mixture of the 2 can be expected to have a potency about 80 % of the undiluted sediment, because the diluent sediment contributes comparatively little clean organic carbon despite the much larger volume and mass dilution. The effect would be reversed if a contaminated sediment low in organic carbon were diluted with a clean sediment high in organic carbon. Similar effects can be expected when diluting a metal-contaminated sediment with a diluent sediment with different AVS content. Also, diluting a field sediment with an inert diluent like sand can be expected to have little or no effect on toxicity because the sand lacks the constituents that control contaminant bioavailability (for example, Ingersoll et al. 2015 (85)). For practical reasons, clean and uncontaminated

comprising a sediment dilution will be apportioned by either mass or volume; for this reason, dilution calculations need to account for differences in density and dry matter content in addition to concentrations of organic carbon, AVS, or other sediment constituents. Finally, diluting a contaminated sediment with an uncontaminated sediment should be viewed as a form of sediment spiking, and should consider appropriate equilibration times and chemical monitoring as discussed in 10.6. While physical homogenization of the 2 sediments can be achieved quickly, chemical equilibration may require weeks or even months.

10.4.3 Sampling Pore Water:

10.4.3.1 *General*—Pore water, sometimes called interstitial pore water, refers to the water occupying the spaces between sediment or soil particles, and is often isolated for chemical characterization or for toxicity testing. Some sediment assessment approaches, such as equilibrium partitioning theory (EqP; DiToro et al. 1990 (4), 1991 (2); USEPA 2003 (23), 2005 (24)) use concentration of chemicals in the pore water as an index of chemical bioavailability and expected toxicity. Toxicity tests using pore water are sometimes used as part of procedures to identify the cause of sediment toxicity (USEPA 2007) (6). The usefulness of pore-water sampling will depend on the study objectives and nature of the sediments at the study site.

10.4.3.2 *Sampling Approaches*—Approaches for sampling pore water vary, but generally fall into 2 categories, physical separation or passive sampling (diffusion/equilibrium with an exogenous sampler). The common physical isolation procedures include: (1) centrifugation (Ankley and Schubauer-Berigan 1994) (101), (2) compression/squeezing (Carr and Chapman 1995) (94), and (3) suction/vacuum (Winger and Lasier 1993 (62), Winger et al. 1998 (102)). Passive sampling does not rely on physically isolating pore water, but instead uses separate media to either provide a sample matrix that has the same concentrations of contaminants, or one that can be used to estimate those concentrations. “Peeper” type samplers rely on the movement (diffusion) of pore-water constituents across semipermeable membranes into a water-filled collecting chamber until concentrations approach equilibrium. Other approaches use solid materials, such as pieces of polyethylene (PE) or polyoxymethylene (POM), that sorb or bind contaminants present in the pore water with known or measurable partition coefficients that can be used to infer chemical concentrations in pore water. Solid-phase microextraction (SPME) fibers have also been applied to sorb chemicals from bulk sediment samples (or from porewater samples isolated from sediment) in order to estimate chemical concentrations in pore water. Reviews by Bufflap and Allen (1995a) (103), Guide E1391, and USEPA (2001 (89)) provide detailed discussion of pore-water extraction and sampling techniques.

10.4.3.3 Passive Samplers:

(1) *Peepers*—Peepers are most commonly used for passive sampling of metal concentrations, and sometimes other inorganic ions such as ammonia or sulfide. Brumbaugh et al. (2013) (104) describes the application of peeper samples for

monitoring metal concentrations during whole sediment toxicity tests. A broader discussion of peeper applications for sampling metals in pore water is provided by Peijnenburg et al. (2014) (105).

(2) *Other Passive Samplers for Metals*—While peepers are probably the most frequently used passive samplers for metals, there are other passive sampling approaches for metals in sediment pore water, such as diffusive gradient in thin films (DGT), and diffusive equilibration in thin film (DET) sampling. Peijnenburg et al. (2014) (105) provide a review of these and other approaches in water and sediment; guidance for using DGT samplers in laboratory settings is provided by USEPA/SERDP/ESTCP (2017). (106)

(3) *Passive Sampling of Organic Chemicals*—USEPA/SERDP/ESTCP (2017) (106) describes the use of 3 passive samplers including low density polyethylene (LDPE), polyoxymethylene (POM), and polydimethylsiloxane (PDMS) coated solid-phase microextraction (SPME) fibers for measuring freely-dissolved hydrophobic organic contaminants in pore water. The LDPE and POM are thin film polymers (that is, like plastic bags) while SPME fibers consist of a thin annular layer of sorbent coating on a small-diameter silica core. All 3 samplers can be configured to be inserted into sediments during toxicity testing. Many example applications of these techniques for direct insertion into the sediment have been published (for example, Mayer et al. 2000 (107), Conder et al. 2003 (108), You et al. 2007 (109), Lu et al. 2011 (110), Gschwend et al. 2012 (111), Ingersoll et al. 2014 (100)).

(4) *Physical Separation*—There is no one method for the isolation of pore water that is unilaterally superior. Factors to consider in the selection of an isolation procedure may include: (1) volume of pore water needed, (2) ease of isolation (materials, preparation time, and time required for isolation), and (3) potential for artifacts caused by the isolation procedure. Each approach has unique strengths and limitations (Bufflap and Allen 1995a (103) and b (112), Winger et al. 1998 (102)), which vary with sediment characteristics, chemicals of concern, toxicity test methods, and desired test resolution (that is, data quality objectives). Centrifugation has become the most widely used method and is effective for producing larger volumes of pore water that are often required for sediment toxicity testing. However, centrifugation can be ineffective for sediments with a high percentage sand or larger particles. For suction or compression separation, which uses a filter or a similar surface, there may be changes to the characteristics of the pore water compared with separation using centrifugation (Ankley and Schubauer-Berigan 1994 (101), Horowitz et al. 1992 (113)). Side-port suctioning of sediment cores may be more effective for sandy sediments, though multiple cores may be necessary to collect larger volumes. Peepers do not directly isolate pore water, though they can produce aqueous samples with similar chemical composition and may be an alternative to physical separation for some applications. Most peeper configurations are designed to produce relatively small sample volumes for chemical analysis, though larger peepers have

been used to gather larger samples for toxicity testing (Burton 1992 (28), Sarda and Burton 1995 (96)). See USEPA (2001) (89) and Guide E1391 for more information on isolating pore water.

10.5 Sediment Characterization:

10.5.1 General Characterization—The specific parameters to be measured are linked to the purposes of the study. Generally speaking, organic carbon content, percent dry matter, and particle size distribution (percent sand, silt, and clay, or more detailed analyses) are the primary characteristics that should be measured, particularly for studies involving field sediments where there may be considerable variability across samples. Because ammonia is often elevated in contaminated sediments and can cause toxicity, measuring ammonia and pH in the pore water of sediments is also recommended. Other parameters that may have utility include biological oxygen demand, chemical oxygen demand, cation exchange capacity, Eh, total inorganic carbon, total volatile solids, acid volatile sulfide, and oil and grease (or petroleum hydrocarbons). See USEPA (2001) (89) and Guide E1391 for more information on physical and chemical characterization of sediments.

10.5.2 Chemical Contaminants—The types of chemical contaminants to be measured in sediments depends on the nature of the samples and the purpose of the study. Analytical methods should be referenced to previously documented methods or have data to support their applicability to sediment analysis. Data Quality Objectives (DQOs) for chemical analysis should consider the likely toxicity of chemicals, in determining the analytical sensitivity necessary for the study. In addition, the chemical analysis plan should consider other sediment characteristics that affect the bioavailability of sediment contaminants, such as organic carbon and acid volatile sulfide.

10.5.3 Monitoring Sediment Characteristics Over Time—It may be desirable to monitor concentrations of contaminants (or potentially labile characteristics such as acid volatile sulfide) over time, particularly if sediments are stored for extended periods.

10.5.4 Monitoring Sediment Characteristics During Testing—Sediment characteristics may change during toxicity testing. It is generally desirable to include chemical characterization of test sediments at the start and end of sediment exposures, at a minimum. Higher frequency sampling may be important for chemicals that are likely to degrade or be otherwise lost during testing. Dedicated replicate chambers can be included in sediment tests to be destructively sampled for chemical analysis at specified intervals. Test organisms and food should be added to these extra chambers to help ensure their conditions match as closely as possible those in replicates used for biological measurements (for example, Besser et al. 2013 (83), Ingersoll et al. 2014 (100)). Sediment samples collected after test sediments have been placed into test chambers can be recovered by siphoning most of the overlying water without disturbing the surface of the sediment, then removing appropriate aliquots of the sediment for chemical analysis.

10.6 Sediment Spiking:

10.6.1 General—Spiking sediments with chemicals can be an important tool in contaminated sediment assessment, used most commonly to create exposure-response relationships that can be used to interpret the hazard associated with chemical concentrations measured in field sediments, or in some cases those that might be predicted based on chemical fate models. Spiked sediment tests can also be used to study the influence of sediment physico-chemical characteristics on chemical toxicity (Adams et al. 1985 (1), Simpson et al. 2004 (114), Hutchins et al. 2009 (115), Brumbaugh et al. 2013 (104), Besser et al. 2013 (83)). Other uses might include determining relative species sensitivity, or assessing the biological endpoints most likely to be affected by exposure to a particular chemical.

10.6.2 Overview of Spiking Process and Objectives—Many sediment spiking methods have been used; the methods used for a particular study need to be matched to the characteristics of the chemical and the purposes of the study. The spiking process can be broken down into 2 primary elements, one involving the introduction of chemical to the sediment, and the second involving the subsequent handling of the sediment before testing, often referred to as the equilibration time period. The method of chemical introduction is generally tied to the properties of the chemical; most chemicals with sufficient water solubility are introduced by dissolving them in a small amount of water, while those that are more hydrophobic are typically introduced via organic solvents, with subsequent steps taken to remove the solvent from the system. Equilibration procedures need to be matched to the study purpose. In many cases, contaminants in field sediments have been in place for long periods of time, resulting in something close to thermodynamic equilibrium within the sediment. Such equilibrium can take weeks or even months to achieve after spiking, and sediments that are tested too soon after spiking may not represent the exposure-response relationships that would exist in the field. On the other hand, if chemicals are rapidly degraded or lost from sediment (for example, if they are degraded fairly quickly by microbes in sediment), long equilibration periods may be infeasible, or even inconsistent with the purpose of the study.

10.6.3 General Considerations—This section discusses general guidelines for spiking sediments. Sections 10.6.4 and 10.6.5 provide additional, more specific guidance applicable to spiking particular types of chemicals.

10.6.3.1 Northcott and Jones (2000) (116) Spiking Method Review and Recommendations—Northcott and Jones (2000) (116) provide a review of methods used to spike hydrophobic organic chemicals into sediment or soils and provide a more detailed discussion of the advantages and disadvantages of various spiking techniques, and several aspects of this discussion also apply to spiking inorganic chemicals. Most of the methodological recommendations from Northcott and Jones (2000) (116) are incorporated into the spiking guidance provided here, though there are exceptions. For example, the recommendation to spike sediment within 2 weeks of collection is not included here, as experience has shown sediment can be successfully used in sediment spiking studies even after months or years of storage, as long as the sediment characteristics remain appropriate to the study objectives.

10.6.3.2 Sediment Selection—Sediment spiking is usually carried out on control sediments or reference sediments, though there may be circumstances where other types of sediments are spiked (including formulated sediments). The sediment should be well characterized (for example, organic carbon, AVS, particle size, background contamination) and should have qualities that are appropriate to the study purpose (for example, organic carbon content). In general, spiking inert substrates like sand is not likely to yield information relevant to natural sediments because they lack the binding phases that control chemical bioavailability in natural sediments.

10.6.3.3 Test Material—The test material(s) should be at least reagent grade, unless a test using a formulated commercial product, technical-grade, or use-grade material is specifically needed. Characteristics of the test material that may be important to designing a spiked sediment study include: (1) the identity and concentration of major ingredients and impurities, (2) water solubility, (3) log Kow (octanol-water partitioning coefficient), BCF (bioconcentration factors from other test species), persistence, hydrolysis, and photolysis rates, (4) toxicity, or estimated toxicity, to the test organism, other aquatic organism, and to humans, (5) if the test concentration(s) are to be measured, the sensitivity, precision, and bias of the analytical method, and (6) recommended handling and disposal procedures.

10.6.3.4 Selecting Spike Concentrations—Chemical concentrations might be normalized to factors other than dry weight. For example, concentrations of nonpolar organic compounds might be normalized to sediment organic-carbon content, and metals might be normalized to acid volatile sulfides (for example, Di Toro et al. 1990 (4), 1991 (2); USEPA 2003 (23), 2005 (24); Burgess et al. 2013 (117)). If a test is intended to generate an exposure-response curve to estimate effect levels (for example, an LC50 or an EC20), a series of spiked concentrations should be selected that will provide partial effects, preferably at 2 or more concentrations of the test chemical. This is often done by developing a logarithmic concentration series spaced by a common factor such as 0.5× (that is, each concentration in the series is one half of the next higher). Effect concentrations for a particular compound may vary depending on physical and chemical characteristics of the sediment. It may be desirable to conduct a range-finding test with widely spaced concentration to narrow the range of concentrations to be tested in a later definitive test (for example, testing 3 or more concentrations of the test material spaced by a factor of 10. Results from water-only toxicity tests could be used to establish concentrations to be tested in a whole-sediment test based on predicted pore-water concentrations (Di Toro et al. 1990 (4), 1991 (2)). In some situations a study may be intended to simply determine whether a specific concentration of test material is toxic to the test organism, and therefore involve only a single concentration, or whether adverse effects occur above or below a specific concentration.

10.6.3.5 Procedural or Spiking Controls—It is important that the study design for spiked sediment testing address the potential for artifacts introduced by the spiking method. For water column toxicity tests that introduce a chemical through carrier solvent or means other than direct dissolution of

chemical into water, the inclusion of a solvent control, in addition to a negative control, is standard practice. The purpose of the solvent control is to determine whether the process used to introduce the test material exerts adverse effects separate from the test material. For spiked sediment testing, the same concepts generally apply, but the definition of appropriate controls may be more complex. If the sediment used for spiking is the same as the laboratory control sediment, then a procedural control can be prepared by passing an aliquot of control sediment through the same processing and equilibration as is used for spiked sediment(s), but with no test material added. If the amount of solvent or solution is not consistent across treatments, the procedural blank should contain the highest amount used in any treatment and should be from the same batch used to make the stock solution (Guide E729, Guide E1241). If the sediment being spiked is not the same as the laboratory control, then 2 types of additional controls may be relevant: (1) an aliquot of the control sediment to be spiked that has not be manipulated at all; and (2) an aliquot of the control sediment processed without added chemical as a procedural control. Another complicating factor is that for water column testing, the goal is to have spiking procedures that perform equivalently to the negative control. Because of the greater complexity of sediments as a spiking and exposure matrix, this may not always be the case; it may not be feasible or possible to have an effective spiking and equilibration procedure that has no effect (positive or negative) on organism performance. Thus, the inclusion and interpretation of performance in various controls may be study specific. All negative controls and performance controls should meet control performance criteria applicable to the test method (Annex A1 to Annex A8). If a difference exists among the controls, the procedural blank is generally the appropriate comparison for evaluating test results. If no statistically significant difference is detected among controls, the combined data from the controls should be used as the basis for calculating the results. If performance in the solvent control is markedly reduced from that in the unmanipulated sediment, it is possible that the data are compromised by experimental artifacts and may not accurately reflect the toxicity of the chemical in natural sediments.

10.6.3.6 Methods for Adding Test Material to Sediment for Chemicals with High Water Solubility—The primary factor affecting the method used for adding chemical to sediment is water solubility. If chemicals are sufficiently soluble in water (for example, with a log Kow below about 3), the chemicals can be added directly to sediment, or dissolved in small aliquot of water (for example, a volume 20 % or less of the native water content of the sediment). For ionizable organic compounds, or inorganic compounds that act as acids or bases, it may be appropriate to adjust the pH of the spiking solution before adding it to the sediment, to avoid disrupting sediment chemistry. Additional considerations applicable to spiking water soluble compounds are discussed in 10.6.4 on spiking metals.

10.6.3.7 Methods for Adding Test material to Sediment for Chemicals with Lower Water Solubility—For compounds that are not highly soluble in water, primarily organic compounds

compounds with a log K_{ow} above about 3 alternative spiking methods must be used. While many of these involve the use of organic solvents, methods that involve direct addition of organic solvent to sediment should be avoided wherever possible, as the addition of solvent can alter sediment chemistry (Northcott and Jones 2000) (116). More specific methods applicable to chemicals with lower water solubility are described in 10.6.5.

10.6.3.8 Initial Mixing—Except for chemicals added by shell coating (10.6.5.4), sediment is mixed immediately after the addition of test material(s). This may be accomplished by several methods, such as a rolling mill, feed mixer, stainless steel auger or paint mixer, or hand mixing with a stainless steel spoon or paddle. The goal of initial mixing is to distribute the chemical evenly throughout the sediment to minimize diffusion distances as the spiked chemical interacts with sediment constituents. As prolonged mixing can alter the texture and structure of sediment, sediment should not be mixed longer or more aggressively than is necessary to distribute a chemical and aid equilibration (usually minutes to hours); mixing should not be so vigorous as to entrain air in the sediment. Rather than continuous mixing, intermittent mixing may be as effective and reduce the overall disturbance of the sediment. Mixing of sediment periodically during equilibration may also be needed (for example, weekly; 10.6.3.9).

10.6.3.9 Equilibration—For most sediment spiking studies, the goal is to assess the potency of the chemical as it would exist in field sediments where chemicals have been in place for long periods and have reached an equilibrium distribution among the various components and binding phases within the sediment (10.6.3.12 for discussion of non-equilibrium conditions). Accordingly, after initial mixing, spiked sediments are generally held in storage, with or without periodic mixing, to allow equilibration. The equilibration temperature should be kept low (for example, 4 °C) to minimize potential changes in the physico-chemical and microbial characteristics of the sediment (Guide E1391). Because it does not require that storage vessels be opened, a roller mill can be an effective means of intermittently remixing sediments, but sediments should not be rolled continuously as this can alter the texture and structure of sediments (Northcott and Jones 2000) (116). Some sediments will settle and stratify to an extent that rolling does not provide sufficient energy to remix the sediments, so sediments should be checked to ensure that they are being mixed adequately. Other mixing methods can be used during equilibration, though the method and frequency should be chosen so as to not unduly affect the spiked sediment (for example, through loss of volatile chemical, or introduction of oxygen).

10.6.3.10 Monitoring Equilibration—Periodic monitoring of chemical concentrations in pore water during sediment aging is recommended as a means to assess the equilibration of the spiked sediments. In most cases, the concentration of spiked chemical in the pore water will start high and decline to a stable concentration over time.

10.6.3.11 Duration of Equilibration—For organic chemicals, important factors affecting the time to achieve equilibrium are the chemical's water solubility and organic

carbon partition coefficient (K_{OC}), and the spiking concentration. Water solubility is important because in most situations, the pore water is the medium through which chemical is delivered to sorption sites on the solid phase, so the larger the capacity, the faster the equilibration. Partition coefficient is important because the larger the K_{OC}, the greater the proportion of the spiked chemical that must be transported to sorption sites. Spiking concentration affects equilibration because higher spiking concentrations require more chemical to be redistributed through the sediment. Organic chemicals with log K_{OW} <4 may reach equilibrium in 1 to 2 weeks, while those with log K_{OW} >6 can take 2 months or more (USEPA Duluth laboratory, unpublished data).

10.6.3.12 Testing Spiked Sediments that are not Fully Equilibrated—In some cases, study objectives or other circumstances may dictate that sediments are tested without being fully equilibrated before testing. Examples might include evaluation of temporally variable exposures (for example, such as contaminants entering system through episodic runoff events) or spiking chemicals that degrade at rates that make full equilibration infeasible. Because the parameters for such studies are likely to be study specific, it is not possible to develop broadly applicable guidance for spiking and equilibration. However, because toxicity tests conducted with unequilibrated sediments can be expected to be very dynamic, it is likely that intensive monitoring of chemical concentrations in sediment and pore water during those tests will be required to allow interpretation of the results.

10.6.3.13 Homogeneity—Homogeneity of spiked sediments is important and should be documented through chemical analysis of subsamples from each container, and among multiple containers if all sediment is not equilibrated in the same container. Verifying homogeneity before testing is most important, but sampling at earlier points in the spiking and equilibration process may provide important additional information. If multiple containers of the same sediment spike are used in a study, combining and homogenizing the smaller containers into 1 larger container is recommended to ensure consistency among replicate test chambers. The required degree of homogeneity should be based on the data quality objectives for the study; 20 % is a common default.

10.6.3.14 Documentation of Spiking Methods—Because the spiking methods used can be a substantial influence on the nature and toxicity of spiked sediments, study reports should include detailed descriptions of the characteristics of the sediments spiked, the methods used for spiking, and information on the equilibration methods (including chemical monitoring of equilibration).

10.6.4 Methods for Spiking Cationic Metals:

10.6.4.1 Spiking Using Water Solutions—Metals are generally added in an aqueous solution (Simpson et al. 2004 (114), Hutchins et al. 2009 (115), Brumbaugh et al. 2013 (104), Guide E1391). Ammonia has also been successfully spiked using aqueous solutions (Besser et al. 1998) (118). The volume of water used for spiking should not be excessive, with the intent of having water content just sufficient to allow adequate mixing as wet sediment slurry (for example, 50 to 70 % total water content, depending on particle size). In addition, a chemical

(spiking) blank is recommended, that is, an aliquot of sediment to which the same amount of water is added (without metal), then manipulated and equilibrated in parallel with spiked sediments.

10.6.4.2 Limiting Oxidation—Most fine grained sediments are anoxic a few millimeters below the sediment-water interface, and methods for spiking metals should be designed to minimize oxidation to the extent practicable. Steps that can be taken include using deoxygenated water to prepare spiking solutions, and mixing sediments in at inert atmosphere (for example nitrogen).

10.6.4.3 Equilibration—Achieving adequate equilibration of the added metal is generally the most important consideration in spiking metals into sediments. Equilibration times of weeks to months may be required for pore-water metal concentrations to stabilize, and times differ between metals and sediments (Simpson et al. 2004) (114). Equilibration times may be shorter for sediments with higher concentrations of fine particles and organic matter than in sandy sediments. Equilibration times may be longer in more oxidized sediments than in more sulfidic sediments. The detection of unnaturally high pore-water metal concentrations (that is, well beyond the range expected for sediments naturally contaminated with similar metal concentrations) is generally indicative that equilibrium has not yet been re-established or that spiking levels have exceeded the sediment binding capacity (Simpson et al. 2004 (114), Hutchins et al. 2009 (115), Brumbaugh et al. 2013 (104), Vandegheuchte et al. 2014 (119)).

10.6.4.4 Monitoring and Adjustment of pH—The disturbance of sediments during the spiking process will often result in oxidation of Fe(II) to Fe(III); Fe(II) exists naturally in pore waters and can be further increased by its displacement from sediment particles by added metals. Hydrolysis and precipitation of Fe(III)-hydroxide phases will release H⁺ and often lower the pH of the sediment. Very high dissolved iron concentrations (>100 mg/L) in pore waters can result. In addition, metals spiked at relatively high concentrations can cause pH depression through hydrolysis reactions, that is, complexation of the spiked metal ion with OH⁻ ions originating from water molecules, which in turn releases H⁺ ions. The expected and desired pH of the final test sediment should be known (often the natural pH), and this may differ considerably for different freshwater sediments. If the pH decreases following spiking, it is recommended that pH be adjusted to the desired pH through small incremental additions of a strong base (for example, 1 M NaOH). The pH of metal-spiked sediments should be checked regularly during the equilibration period (for example weekly) as further pH-adjustment may be necessary.

10.6.4.5 Preparing a Concentration Series Via Dilution—While direct spiking to achieve a specific metal concentration can be effective (Simpson et al. 2004) (114), the dilution of a ‘superspike’ (a very high spiked concentration, also referred to as indirect spiking) is recommended as this method generally results in shorter time to achieve the desired level of metal partitioning (low porewater metals; Hutchins et al. 2009 (115), Brumbaugh et al. 2013 (104)).

10.6.4.6 Introduction of Metal Spiked Sediments into Test Systems—After equilibration, sediments spiked with extremely high metal concentrations may exhibit abnormally high metal concentrations in pore water and overlying water (and possibly Na⁺ ions resulting from pH adjustment) simply because the capacity of the sediment to bind the metal has been exceeded. In such instances it is advisable to remove “excess” metal from the overlying water of the sediment before testing. This can be accomplished by decanting off the overlying water from the spiking container or replacing overlying water often in test chambers before testing (Besser et al. 2013) (83). Also, the homogenization process destroys the oxic-anoxic stratigraphy found in bedded sediments. A surficial oxic layer is reformed after the sediment is placed into test chambers and covered with oxic overlying water. Oxidation of the surficial layer of a metal spiked sediment can release metals to the overlying water in a manner that may not be representative of most field sediments. As noted in Annex A1 to Annex A5 allowing pre-test equilibration for more than the default 24-h period may be appropriate for some studies.

10.6.5 Methods for Spiking Non-polar Organic Chemicals:

10.6.5.1 General—Non-polar organic compounds present challenges related primarily to their limited water solubility, and the associated difficulties with introducing the chemical into the sediment in a way that effectively distributes the chemical while avoiding the introduction of organic solvents or other substances that would unacceptably alter sediment characteristics. Methods that provide greater dispersion or surface area of the added chemical can speed equilibration of the chemical.

10.6.5.2 Containers—Because hydrophobic organic chemicals interact with many plastics, safety-coated glass jars may be useful containers for spiking and equilibrating sediments. These jars are sold in a range of sizes, are compatible with most roller mixers, and the anti-shatter safety coating provides some worker protection should a container break.

10.6.5.3 Direct Addition—USEPA (2016a) (120) describes addition of solid chemicals directly to sediment followed by thorough mixing. To aid in dispersion and dissolution, grinding the test material to particles <200 µm is recommended. When adding test material to the sediment, steps should be taken to ensure the particles of test material are well dispersed and do not clump.

10.6.5.4 “Wall” or “Shell” Coating—The shell coating method involves coating the inside of the spiking vessel with a layer of crystallized chemical to provide a dispersed thin layer of chemical for dissolution (Ditsworth et al. 1990) (121). Test material is dissolved in a small amount of volatile organic solvent (for example, acetone, benzene) and introduced to an empty spiking vessel. The vessel is then placed on a roller so the spiking solution is dispersed over the walls of the vessel while the solvent evaporates. Applying a vacuum to the vessel can speed evaporation of the solvent; gentle heating can also be used, provided it does not cause the test material to evaporate or degrade. It is important that all solvent is eliminated before introducing the test sediment. Unlike many other methods wherein the sediment is homogenized after the test material is

introduced, spiking via shell coating relies on gradual dissolution of the test material from the walls of vessel into the pore water. As such, mixing via slow (for example, 2 revolutions per minute) rotation on a roller, at intervals over time, is generally the best method of mixing for use with shell coating. However, continuous rolling for long periods should not be used as it will change the texture and structure of sediments.

10.6.5.5 Sand Coating—In the sand coating method, test material is dissolved in organic solvent and combined with a small amount of fine sand, then the solvent is evaporated leaving the test material crystallized on the sand particles (USEPA 2019) (7). This not only provides a large surface area, but also allows the crystals to be easily dispersed throughout the sediment, resulting in smaller diffusion distances and less need for frequent mixing during equilibration. Relatively fine sand is preferred to maximize surface area; the amount of sand used is flexible, though generally kept to a small percentage of the sediment volume to avoid substantially altering the sediment characteristics. Sand coating can be combined with the shell coating technique by adding the solvent-wetted sand to the spiking vessel, then evaporating the solvent as described for shell coating. Sand coating can also be used for test materials that are liquid at test or incubation temperatures. After the solvent is evaporated, the coated sand and sediment are combined and thoroughly mixed.

10.6.5.6 Addition to Dried Sediment—Some spiked sediment studies have been conducted by combining test material in organic solvent with a portion of sediment that has been dried (for example, as in Ingersoll et al. 2005a) (122). While this avoids the introduction of sand that was not originally part of the sediment. However, the process of drying sediment can be expected to alter the chemistry and characteristics of the sediment, so this approach is not recommended unless there is a specific reason to favor it over other alternatives (Northcott and Jones 2000) (116).

10.6.5.7 Direct Addition of Solvent or Surfactant and Dispersant to Wet Sediment—Some spiked sediment studies have been conducted by adding a solvent and test material solution directly to wet sediment. This is not recommended because of the expected effects on the sediment and porewater composition and the potential for increasing biological oxygen demand (Northcott and Jones 2000) (116). Likewise, surfactants or dispersants should not be used, not only because of effects on sediment chemistry, but also because of their likely influence on chemical bioavailability. An exception might be when a test material includes a surfactant or dispersant as part of its formulation (for example, formulated pesticides; USEPA 2016a) (120).

10.6.5.8 Equilibration—As discussed in 10.6.3.11, the period required for equilibration will depend on the spiked concentration, water solubility of the chemical, and the partition coefficient. Periodic mixing (but not continuous) during equilibration is recommended, particularly for the shell coating method wherein chemical is not initially dispersed throughout the sediment. Chemical concentrations in pore water are typically used to monitor equilibration.

11. Quality Assurance and Quality Control

11.1 Introduction:

11.1.1 Elements of a Quality Assurance Program—Developing and maintaining a laboratory quality assurance (QA) program requires an ongoing commitment by laboratory management and also includes the following: (1) appointment of a laboratory quality assurance officer with the responsibility and authority to develop and maintain a QA program, (2) preparation of a Quality Assurance Project Plan (QAPP) with Data Quality Objectives (DQOs), (3) preparation of written descriptions of laboratory Standard Operating Procedures (SOPs) for test organism culturing, testing, instrument calibration, sample chain-of-custody, laboratory sample tracking system, and (4) provision of adequate, qualified technical staff and suitable space and equipment to assure reliable data. Additional guidance for QA can be obtained in USEPA (1989) (123) and in USEPA (1994e) (124).

11.1.2 Coverage—QA practices within a testing laboratory should address all activities that affect the quality of the final data, such as: (1) sediment sampling and handling, (2) the source and condition of the test organisms, (3) condition and operation of equipment, (4) test conditions, (5) instrument calibration, (6) replication, (7) use of reference toxicants, (8) record keeping, and (9) data evaluation.

11.1.3 Quality Control—Quality control (QC) practices consist of the more focused, routine, day-to-day activities carried out within the scope of the overall QA program. For more detailed discussion of quality assurance, and general guidance on good laboratory practices related to testing see FDA (1978) (125), USEPA (1979a) (126), USEPA (1980a) (127), USEPA (1980b) (128), USEPA (1991a) (55), USEPA (1994c) (129), USEPA (1994d) (64), USEPA (1995) (130), DeWoskin (1984) (131), and Taylor (1987) (132).

11.2 Performance-based Criteria:

11.2.1 Background and Definition—The USEPA Environmental Monitoring Management Council (EMMC) recommended the use of performance-based methods in developing standards for chemical analytical methods (Williams 1993) (133). Performance-based methods were defined by EMMC as a monitoring approach that permits the use of appropriate methods that meet pre-established demonstrated performance standards. Minimum required elements of performance, such as precision, reproducibility, bias, sensitivity, and detection limits should be specified, and the method should be demonstrated to meet the performance standards.

11.2.2 Performance-based Criteria Application to Organism Culture—Participants at a September 1992 USEPA sediment toxicity workshop arrived at a consensus on several culturing and testing methods for freshwater organisms (Appendix A of USEPA 1994a (134)). Experience since that time has indicated that a variety of culture methods can produce healthy organisms for use in sediment toxicity tests, and that no single method is so uniquely superior that it needs to be used to culture test organisms. While having healthy organisms of known age and consistent quality is the key consideration in selecting and developing culture methods, the adequacy of those methods is primarily a function of being able to produce organisms in of sufficient quality and health that the test organisms will consistently exceed control performance criteria when used for testing. Those control performance criteria

are outlined as part of each of the test methods described in this standard (Table A1.1, Table A2.1, Table A3.1, Table A4.1, Table A5.1, Table A6.1, Table A7.1, Table A8.1).

11.2.3 Performance-based Criteria Application to Individual Tests—Performance-based methods were also established to determine the acceptability of data generated from individual sediment (or water) toxicity tests. Table A1.1, Table A2.1, Table A3.1, Table A4.1, Table A5.1, Table A6.1, Table A7.1, and Table A8.1 provide a listing of test acceptability criteria (TAC) for control performance of individual tests. The TAC present a minimum level of performance to judge the quality of data from individual tests. An additional level of performance-based evaluation is provided by laboratory proficiency guidelines as outlined in 11.2.4 and 11.9.

11.2.4 Performance-based criteria for Laboratory Proficiency—The TAC for control performance represent only minimally successful organism performance for single tests; the TAC should represent the lower end of the range of organism performance, not the norm. Across multiple tests, higher control performance should be expected from laboratories proficient with the test methods. If control performance typical for a laboratory is not substantially above minimum TAC, it may be indicative of underlying problems with laboratory facilities, practices, or staff experience. Laboratory proficiency guidelines represent a “performance-based” means of documenting the ability of a laboratory to successfully execute a sediment toxicity test method over time. Laboratory proficiency guidelines of individual test methods are provided in Table A1.1, Table A2.1, Table A3.1, Table A4.1, and Table A5.1. It is important to emphasize that laboratory proficiency guidelines are not intended to be used to judge the validity of individual tests, only to assess performance over time. Plotting performance in control media over time against these guidelines may help laboratories identify trends in control performance.

11.3 Facilities, Equipment, and Test Chambers:

11.3.1 Preventing Cross-contamination—Areas used for test organism culturing and testing should be designed and operated to minimize the potential of cross-contamination between cultures and sediment tests. Ventilation systems should be designed and operated to prevent recirculation or leakage of air from chemical analysis laboratories or sample storage and preparation areas into test organism culturing or sediment or water testing areas, and from sediment or water testing laboratories and sample preparation areas into culture areas.

11.3.2 Temperature Control—Equipment for temperature control should be adequate to maintain recommended test-water temperatures. Recommended materials should be used in the fabricating of the test equipment that comes in contact with the sediment or test water.

11.3.3 Demonstrating Adequacy through Control Tests—Before a sediment or water test is conducted in a new facility, a “control” test should be conducted in which all test chambers contain a control sediment or control sand and overlying water. This information is used to demonstrate that the facility, staff, test procedures, control sediment or control sand, food, and water provide acceptable responses of test organisms (11.9).

11.4 Water—The quality of water used for organism culturing and testing is important. Overlying water used in testing and water used in culturing organisms should be uniform in quality. Acceptable water should allow satisfactory survival, growth, or reproduction of the test organisms. Test organisms should not show signs of disease or apparent stress (for example, discoloration, unusual behavior). In particular, waters used to culture or test *H. azteca* should have a minimum concentration of about 0.02 to 0.04 mg Br/L. See 7.1 for additional details.

11.5 Sediment Collection, Storage, Characterization, and Spiking—Sediment collection, storage, and manipulations should be consistent with the methods and considerations described in Section 10.

11.6 Test Organisms:

11.6.1 Source and Identification—All organisms in a test must be from the same culture cohort (11.2.2). A cohort of test organisms may be obtained from laboratory cultures or from an outside source, provided the outside source can provide documentation that the cohort of organisms meets applicable requirements. The species of test organisms should be positively identified using an appropriate taxonomic key, and verification should be documented (e.g., Pennak 1989 (135), Merritt and Cummins 1996 (136)). Obtaining organisms from wild populations should be avoided unless organisms are cultured through several generations in the laboratory (perhaps a minimum of 2 generations). The ability of the wild population of sexually reproducing organisms to cross breed with the existing laboratory population should be determined (Duan et al. 1997) (137). The sensitivity of the wild population to select chemicals should also be documented.

11.6.2 Special Considerations for *Hyalella azteca*—Organisms morphologically identified as *H. azteca* belong to a genetically and geographically diverse species complex (Major et al. 2013) (49). In a genetic study of *H. azteca* collected from 15 aquatic toxicology laboratories, Major et al. (2013) (49) found that all but 1 had organisms that were genetically similar, belonging to a group they called the “US Laboratory clade” (though its distribution included several Canadian laboratories as well). A single laboratory, the Canadian Centre for Inland Waters in Burlington, ON, had *H. azteca* of another clade. The research underpinning the test methods for *H. azteca* was conducted almost exclusively using organisms known or thought to be of the US Laboratory clade and, as such, the methods and specifications are believed to be most applicable to amphipods of that clade. Limited comparative work across clades in the laboratory (Soucek et al. 2013) (58) and more extensive study with field populations (Strong 1972 (138), Wellborn 1993 (139) indicates there are differences in size, fecundity and perhaps physiology among clades. Researchers applying the methods described here using different clades should be aware of the potential for differences in appropriateness of test methods and comparability of results with studies performed with the US Laboratory clade.

11.6.3 Health of Test Cohort—The cohort of organisms used for any test should appear healthy, behave normally, feed well, and have low mortality in cultures, and during pre-test holding (for example, <10 % for 48 h before the start of a test).

11.7 *Quality of Food*—Problems with the nutritional suitability of the food will be reflected in the survival, growth, or reproduction of the test organisms in cultures or in sediment tests. Periodic monitoring of contaminant concentrations is desirable. Components of food preparations should not be used beyond expiration dates provided by the manufacturers; food preparations should be dated when prepared and not used beyond the recommendations included in the preparation instructions (12.5).

11.8 Test Conditions:

11.8.1 *General*—Conditions for a toxicity test should conform with conditions outlined in Table A1.1, Table A2.1, Table A3.1, Table A4.1, Table A5.1, Table A6.1, Table A7.1, and Table A8.1. Temperature, dissolved oxygen, alkalinity, water hardness, conductivity, ammonia, and pH should be checked as described for each test. In addition, it is desirable to have equipment to monitor temperature continuously in the testing apparatus (for example, in a test chamber or the surrounding water bath) to provide assurance that temperature excursions are not occurring outside observation periods. Connection of temperature monitoring equipment to an alarm system that can notify staff of temperature excursions is recommended.

11.8.2 *Acceptability*—An individual test may be conditionally acceptable if temperature, dissolved oxygen, and other specified conditions fall outside specifications, depending on the degree of the departure and the objectives of the tests. The acceptability of a test will depend on the experience and professional judgment of the laboratory analyst and the reviewing staff of the regulatory authority. Any deviation from test specifications should be noted when reporting data from a test.

11.9 Preliminary Testing with Negative Controls:

11.9.1 *General*—Before conducting a sediment or water test, it is strongly recommended that the laboratory conduct tests with control sediment(s) or control sand alone. The ability to meet both test acceptability criteria (TAC; 11.2.3) in single tests and laboratory proficiency criteria (11.2.4) over multiple tests is an important demonstration of the aggregate adequacy of facilities, procedures, organisms, and staff. Preliminary testing with negative control media can also be used to diagnose problems and avoid costly failures of later tests.

11.9.2 *Organism Recovery*—An important aspect of sediment toxicity testing is the successful recovery of organisms from sediment. Laboratories should demonstrate that their personnel are able to recover an average of at least 90 % of the organisms from whole sediment. For example, random numbers of test organisms could be added to test chambers containing control sediment by 1 technician, then a second technician could process the sediment as would be done in a sediment toxicity test to evaluate proficiency in organism recovery (for example, Tomasovic et al. 1994 (140)). Very young amphipods (<7-d old), as might be encountered during life-cycle testing with *H. azteca*, are particularly challenging to recover from sediment, and may be useful as a target for practicing recovery of organisms from sediment (A2.3.4.1(2)).

11.10 Reference Toxicant Testing:

11.10.1 *Background and Intent*—Historically, reference-toxicity testing has been thought to provide 3 types of

information relevant to the interpretation of toxicity test data: (1) an indication of the relative “health” of the organisms used in the test; (2) a demonstration that the laboratory can perform the test procedure in a reproducible manner; and (3) information to indicate whether the sensitivity of the particular strain or population in use at a laboratory is comparable to those in use in other facilities. With regard to the first type of information, McNulty et al. (1999) (141) suggest that reference-toxicity tests are seldom effective in identifying stressed populations of test organisms. In addition, reference-toxicity tests recommended for use with sediment toxicity tests are short-term, water column tests, owing in part to the lack of a standard sediment for reference-toxicity testing. Because the test procedures for reference-toxicity tests are not the same as for the sediment toxicity tests of interest, the applicability of reference-toxicity tests to demonstrate ability to reproducibly perform the sediment test procedures is greatly reduced. Particularly for the long-term sediment toxicity tests with *H. azteca* and *C. dilutus* (Annex A3 and Annex A4), performance of control organisms over time may be a better indicator of success in handling and testing these organisms (11.2.3, 11.2.4, and 11.9). For these reasons, requirements for reference toxicant testing were relaxed in the previous versions of this standard.

11.10.2 *Recommended Frequency*—Although the requirement for monthly testing was removed in this standard, periodic reference-toxicity testing is still recommended as an indication of overall comparability of results among laboratories (for example, 6 tests over a 3-year period). In particular, reference-toxicity tests should be performed more frequently when organisms are obtained from outside sources, when there are changes in culture practices, or when brood stock from an outside source is incorporated into a laboratory culture.

11.10.3 *Common Reference Toxicants*—Reference toxicants such as sodium chloride (NaCl), potassium chloride (KCl), cadmium chloride (CdCl₂), and copper sulfate (CuSO₄) are suitable for use. No one reference toxicant can be used to measure the sensitivity of test organisms with respect to another toxicant with a different mode of action (Lee 1980) (142). However, it may be unrealistic to test more than 1 or 2 reference toxicants routinely. KCl has been used successfully in round-robin water-only exposures with *H. azteca* and *C. dilutus* (14.4.1.1).

11.10.4 *Test Methods*—Test conditions for conducting reference-toxicity tests with *H. azteca* and *C. dilutus* are outlined in Table 7 and Table 8. Reference-toxicity tests can be conducted using one organism/chamber or multiple organisms in each chamber. Some laboratories have observed low control survival when more than 1 midge/chamber is tested in water-only exposures.

11.10.5 *Intra-laboratory Precision*—Intra-laboratory precision, expressed as a coefficient of variation (CV) of the range in response for each type of test to be used in a laboratory, can be determined by performing 5 or more tests with different batches of test organisms using the same reference toxicant at the same concentrations with the same test conditions (for example, the same test duration, type of water, age of test organisms, feeding) and the same data analysis

TABLE 7 Recommended Test Conditions for Conducting Reference-toxicity Tests with 1 Organism/Chamber

Parameter	Specification
Test type	Water-only acute test
Dilution series	Control and at least 5 concentrations spaced by no more than 0.5x
Toxicant	NaCl, KCl, Cd (salt), or Cu (salt); suggested toxicants, others are acceptable
Temperature	23 ± 1 °C
Light quality	Wide-spectrum fluorescent lights
Illuminance	About 100 to 1000 lux
Photoperiod	16h light:8h dark
Renewal of water	None
Age of organisms	<i>H. azteca</i> : 7-10 d old (not more than 2-d range in age) <i>C. dilutus</i> : Around 2 nd instar; consistent with specifications in A3.2.7
Exposure chamber	30-mL plastic cups (covered with glass or plastic)
Water volume	20 mL
Number of organisms/chamber	1
Number of replicate chambers/treatment	10 minimum
Feeding	<i>H. azteca</i> : 0.1 mL YCT on Day 0 and 2 <i>C. dilutus</i> : 0.25 mL fish flake suspension (4 g/L) on Day 0 and 2
Substrate	<i>H. azteca</i> : 1 mL sand, or nitex screen (110 mesh) <i>C. dilutus</i> : 1 mL sand
Aeration	None
Dilution water	Culture water or water used for sediment toxicity testing; for <i>H. azteca</i> , ≥0.04 mg Br/L and ≥15 mg Cl/L
Test chamber cleaning	None
Water quality	Hardness, alkalinity, and conductivity at the beginning of a test; pH and DO at the beginning and end of a test; temperature daily
Test duration	96 h
Endpoint	Survival (LC50)
Test acceptability	90 % control survival

TABLE 8 Recommended Test Conditions for Conducting Reference-toxicity Tests with More Than 1 Organism/Chamber

Parameter	Specification
Test type	Water-only acute test
Dilution series	Control and at least 5 concentrations spaced by no more than 0.5x
Toxicant	NaCl, KCl, Cd (salt), or Cu (salt); suggested toxicants, others are acceptable
Temperature	23 ± 1 °C
Light quality	Wide-spectrum fluorescent lights
Illuminance	About 100 to 1000 lux
Photoperiod	16h light:8h dark
Renewal of water	None
Age of organisms	<i>H. azteca</i> : 7-10 d old (not more than 2-d range in age) <i>C. dilutus</i> : Around 2 nd instar; consistent with specifications in A3.2.7
Exposure chamber	250-mL glass beaker (covered with glass or plastic)
Water volume	100 mL
Number of organisms/chamber	5 minimum (need at least 10 organisms total per treatment)
Number of replicate chambers/treatment	2 minimum
Feeding	<i>H. azteca</i> : 0.5 mL YCT on Day 0 and 2 <i>C. dilutus</i> : 1.25 mL fish flake suspension (4 g/L) on Day 0 and 2
Substrate	<i>H. azteca</i> : 5 mL sand, or nitex screen (110 mesh) <i>C. dilutus</i> : 5 mL sand
Aeration	None
Dilution water	Culture water or water used for sediment toxicity testing; for <i>H. azteca</i> , ≥0.04 mg Br/L and ≥15 mg Cl/L
Test chamber cleaning	None
Water quality	Hardness, alkalinity, and conductivity at the beginning of a test; pH and DO at the beginning and end of a test; temperature daily
Test duration	96 h
Endpoint	Survival (LC50)
Test acceptability	90 % control survival

methods. This serves as a point of reference for future testing. A series of reference toxicant concentrations should be selected that will provide partial mortalities at 2 or more concentrations of the test chemical. To provide sensitivity to changes in sensitivity over time, concentrations in the series should not be more than 0.5x apart; even closer spacing may be necessary to achieve partial effects, depending on the response slope for the reference toxicant used. Information from previous tests can be used to improve the design of subsequent tests to optimize the dilution series selected for future testing.

11.10.6 Interpretation—Outliers, which are data falling outside the control limits, or trends of increasing or decreasing sensitivity can be identified by plotting reference toxicity test results over time. If the reference-toxicity results from a given test fall outside the “expected” range (for example, +2 SD), the sensitivity of the organisms and the credibility of the test results may be suspect. In this case, the test procedure should be examined for defects and should be repeated with a different batch of test organisms (Section 13).

11.10.7 Acceptability—A sediment or water test may be acceptable if specified conditions of a reference-toxicity test fall outside the expected ranges (11.10.2). Specifically, a sediment test should not necessarily be judged unacceptable if the LC50 if the associated reference-toxicity test falls outside the expected range or if mortality in the control of the reference-toxicity test exceeds 10 %. Further, consistency of reference toxicant test results often improves with staff

experience, narrowing the control limits. Control limits of +2 SD, by definition, are expected to be exceeded 5 % of the time, regardless of how well a laboratory performs. For this reason, good laboratories that develop very narrow control limits may be penalized if a test result that falls just outside the control limits is rejected *de facto*. The width of the control limits should be considered in decisions regarding acceptance or rejection of data. All the performance criteria outlined in [Table A1.1](#), [Table A2.1](#), [Table A3.1](#), [Table A4.1](#), [Table A5.1](#), [Table A6.1](#), [Table A7.1](#), and [Table A8.1](#) should be considered when determining the acceptability of a sediment or water test, and will depend on the experience and judgment of the investigator and, if applicable, the regulatory authority.

11.11 Replication and Test Sensitivity—The sensitivity of sediment or water tests will depend in part on the number of replicates/treatment, the significance level selected, and the type of statistical analysis. If the variability remains constant, the sensitivity of a test will increase as the number of replicates is increased. The minimum recommended number of replicates varies with the objectives of the test and the statistical method used for analysis of the data (Section 13).

11.12 Analytical Methods:

11.12.1 Physical and Chemical Analyses—All routine chemical and physical analyses for culture media testing must,

food, and sediment should include established quality assurance practices outlined in USEPA methods manuals (USEPA 1979a and b (126, 74), USEPA 2002d (21)).

11.12.2 *Reagents*—Reagent containers should be dated when received from the supplier, and the shelf life of the reagent should not be exceeded. Working solutions should be dated when prepared and the recommended shelf life should not be exceeded.

11.13 Calibration and Standardization:

11.13.1 *Calibration*—Instruments used for routine measurements of chemical and physical characteristics such as pH, dissolved oxygen, temperature, and conductivity should be calibrated before use each day according to the instrument manufacturer's procedures as indicated in the general section on quality assurance (see USEPA Methods 150.1, 360.1, 170.1, and 120.1; USEPA 1979b (74)). Calibration data should be recorded in a permanent log.

11.13.2 *Standards and Reference Materials*—A known-quality water should be included in the analyses of each batch of water samples (for example, water hardness, alkalinity, conductivity). It is desirable to include certified standards in the analysis of water samples.

12. Collection, Culturing, and Maintaining the amphipod *Hyalella azteca* and the midge *Chironomus dilutus*

12.1 *Life Histories*—The following sections describe life histories and culturing procedures for *Hyalella azteca* and *Chironomus dilutus*. Annex A5 to Annex A8 provide life histories and culturing procedures for additional test organisms.

12.1.1 *Hyalella azteca*:

12.1.1.1 *Occurrence in Nature*—*Hyalella azteca* inhabit permanent lakes, ponds, and streams throughout North and South America (de March 1981 (143), Pennak 1989 (135)). Occurrence of *H. azteca* is most common in warm (20 °C to 30 °C for much of the summer) mesotrophic or eutrophic lakes that support aquatic plants. These amphipods are also found in ponds, sloughs, marshes, rivers, ditches, streams, and springs, but in lower numbers. *Hyalella azteca* have achieved densities of >10,000/m² in preferred habitats (de March 1981) (143). *Hyalella azteca* are epibenthic detritivores that burrow into the sediment surface. Hargrave (1970) (144) reported that *H. azteca* selectively ingest bacteria and algae.

12.1.1.2 *Reproductive Biology*—Reproduction by *H. azteca* is sexual. The adult males are larger than females and have larger second gnathopods (de March 1981) (143). Males pair with females by grasping the females (amplexus) with their gnathopods while on the backs of the females. After feeding together for 1 to 7 d the female is ready to molt and the 2 organisms separate for a short time while the female sheds her old exoskeleton. Once the exoskeleton is shed, the 2 organisms reunite and copulation occurs. The male places sperm near the marsupium of the female and her pleopods sweep the sperm into the marsupium. The organisms again separate and the female releases eggs from her oviducts into the marsupium where they are fertilized. *Hyalella azteca* average about 18 eggs/brood (Pennak 1989) (135) with larger organisms having more eggs (Cooper 1965) (145). The developing embryos and

newly hatched young are kept in the marsupium until the next molt. At 24 °C to 28 °C, hatching ranges from 5 to 10 d after fertilization (Embry 1911 (146), Bovee 1950 (147), Cooper 1965 (145)). The time between molts for females is 7 to 8 d at 26 °C to 28 °C (Bovee 1950). Therefore, about the time embryos hatch, the female molts and releases the young. *Hyalella azteca* average 15 broods in 152 d (Pennak 1989) (135). Pairing of the sexes is simultaneous with embryo incubation of the previous brood in the marsupium. *Hyalella azteca* have a minimum of 9 instars (Geisler 1944) (148). There are 5 to 8 pre-reproductive instars (Cooper 1965) (145) and an indefinite number of post-reproductive instars. The first 5 instars form the juvenile stage of development, instar stages 6 and 7 form the adolescent stage when sexes can be differentiated, instar stage 8 is the nuptial stage, and all later instars are the adult stages of development (Pennak 1989) (135).

12.1.1.3 *Taxonomy of H. azteca*—*Hyalella azteca* belong to a genetically and geographically diverse species complex (Major et al. 2013) (49). In a genetic study of *H. azteca* collected from 15 aquatic toxicology laboratories, Major et al. (2013) (49) found that all but 1 of the laboratories had organisms that were genetically similar, belonging to a group they called the “US Laboratory clade” (though its distribution included several Canadian laboratories as well). The original source of the US Laboratory clade was collected in Oregon by Alan Nebeker of the USEPA in Corvallis OR in the early 1980s. A single laboratory, the Canadian Centre for Inland Waters in Burlington, ON, had *H. azteca* of another clade. The research underpinning the test methods for *H. azteca* was conducted almost exclusively using organisms known or thought to be of the US Laboratory clade and, as such, the methods and specifications are believed to be most applicable to amphipods of that clade. Limited comparative work across clades in the laboratory (Soucek et al. 2013) (58) and more extensive study with field populations (Strong 1972 (138), Wellborn 1993 (139)) indicates there are differences in size, fecundity and perhaps physiology among clades. Culture methods for *H. azteca* described in 12.3 were developed using the US Laboratory clade.

12.1.1.4 *Temperature*—Temperatures tolerated by *H. azteca* range from 0 to 33 °C (Embry 1911 (146), Bovee 1949 (149), Sprague 1963 (150)). At temperatures less than 10 °C the organisms rest and are immobile (de March 1977 (151), de March 1978 (152)). At temperatures of 10 °C to 18 °C, reproduction can occur. Juveniles grow more slowly at colder temperatures and become larger adults. Smaller adults with higher reproduction are typical when organisms are grown at 18 °C to 28 °C. The highest rates of reproduction occur at 26 °C to 28 °C (de March 1978) (152) while lethality occurs at 33 °C to 37 °C (Bovee 1949 (149), Sprague 1963 (150)).

12.1.1.5 *Water Quality*—*Hyalella azteca* are found in waters of widely varying types. *Hyalella azteca* can inhabit saline waters up to 29‰; however, their distribution in these saline waters has been correlated to water hardness (Ingersoll et al. 1992) (65). *Hyalella azteca* inhabit water with high Mg concentrations at conductivities up to 22 000 µS/cm, but only up to 12 000 µS/cm in Na-dominated waters (Ingersoll et al. 1992) (65).

1992) (65). De March (1981) (143) reported *H. azteca* were not collected from locations where calcium was less than 7 mg/L. *Hyaella azteca* have been cultured in reconstituted salt water with a salinity up to 15‰ (Ingersoll et al. 1992 (65), Winger and Lasier 1993) (62). In laboratory studies, Sprague (1963) (150) reported a 24-h LC50 for dissolved oxygen at 20 °C of 0.7 mg/L. Pennak and Rosine (1976) (153) reported similar findings. Nebeker et al. (1992) (154) reported 48-h and 30-d LC50s for *H. azteca* of less than 0.3 mg/L dissolved oxygen. Weight and reproduction of *H. azteca* were reduced after 30-d exposure to 1.2 mg/L dissolved oxygen. In 10-d tests similar to those described in Annex A1, Irving et al. (2004) (42) reported poor survival of *H. azteca* at 1.2 mg/L dissolved oxygen, but survival and growth was similar to control at 2.9 mg/L dissolved oxygen. Also in 10-d tests, Mattson et al. (2007) (43) observed no reduction in biomass of *H. azteca* at the lowest tested oxygen concentration of 2.1 mg/L.

12.1.1.6 *Substrate*—*Hyaella azteca* tolerate a wide range of substrates. Ingersoll et al. (1996) (63) reported that *H. azteca* tolerated sediments ranging from more than 90 % silt- and clay-sized particles to 100 % sand-sized particles without detrimental effects on either survival or growth. *Hyaella azteca* tolerated a wide range in grain size and organic matter in 10- to 42-d tests with formulated sediment (Suedel and Rodgers 1994 (35), Ingersoll et al. 1998) (155). Ankley et al. (1994b) (45) evaluated the effects of natural sediment physico-chemical characteristics on the results of 10-d laboratory toxicity tests with *H. azteca* and *C. dilutus*. Tests were conducted with and without the addition of exogenous food. Survival of organisms was decreased in tests without added food. Physico-chemical sediment characteristics including grain size and TOC were not significantly correlated to the response of *H. azteca* in either fed or unfed tests.

12.1.2 *Chironomus dilutus*:

12.1.2.1 *Occurrence in Nature*—*Chironomus dilutus* inhabit North America (Shobanov et al. 1999) (156) and are commonly found in eutrophic ponds and lakes (Flannagan 1971 (157), Driver 1977 (158)). Midge larvae are important in the diet of fish and waterfowl (Sadler 1935 (159), Siegfried 1973 (160), Driver et al. 1974 (161), McLarney et al. 1974 (162)). Larvae of *C. dilutus* usually penetrate a few cm into sediment. In both lotic and lentic habitats with soft bottoms, about 95 % of the chironomid larvae occur in the upper 10 cm of substrates, and very few larvae are found below 40 cm (Townsend et al. 1981) (163). Larvae were found under the following conditions in British Columbia lakes by Topping (1971) (164): particle size <0.15 mm to 2.0 mm, temperature 0 to 23.3 °C, dissolved oxygen 0.22 to 8.23 mg/L, pH 8.0 to 9.2, conductivity 481 to 4,136 µmhos/cm, and sediment organic carbon 1.9 to 15.5 %. Larvae were absent from lakes if hydrogen sulfide concentration in overlying water exceeded 0.3 mg/L. Abundance of larvae was positively correlated with conductivity, pH, amount of food, percentages of particles in the 0.59 to 1.98 mm size range, and concentrations of Na, K, Mg, Cl, SO₄, and dissolved oxygen. Others (for example, Curry 1962 (165), Oliver 1971 (166)) have reported a temperature range of 0 to 35 °C and a pH range of 7 to 10.

12.1.2.2 *Taxonomy*—Previous editions of this standard have used the species name, *Chironomus tentans*. Since the publication of the second edition, taxonomists have clarified that this midge species is actually *Chironomus dilutus* (Shobanov et al. 1999) (156), which is the species name used in this addition. To be clear, the organism being used is the same, it is only its taxonomy that has changed.

12.1.2.3 *Life Cycle*—The life cycle of *C. dilutus* can be divided into 4 stages: (1) an egg stage (aquatic), (2) a larval stage (aquatic) consisting of 4 instars, (3) a pupal stage (aquatic), and (4) an adult stage (flying/terrestrial). Mating behavior has been described by Sadler (1935) (159). Adult males are easily distinguished from females because males have large, plumose antennae and a much thinner abdomen with visible genitalia. The male has paired genital claspers on the posterior tip of the abdomen (Townsend et al. 1981) (163). The adult female weighs about twice as much as the male, with about 30 % of the female weight contributed by the eggs. After mating, adult females oviposit a single transparent, gelatinous egg mass directly into the water. In the laboratory, females that are provided an opportunity to mate generally oviposit a gelatinous egg mass of 500 to 2000 eggs within 24 h after emergence (personal communication, J. Jenson, Badger Technical Services, Duluth, MN) these eggs hatch in about 3 days at 23 °C. Newly hatched larva remain within the gelatinous egg mass for about 24 h before leaving the egg mass and building and building a case (tube). The cases lengthen and enlarge as the larvae grow with the addition of small particles bound together with threads from the mouths of larvae (Sadler 1935) (159). The larvae draw food particles inside the tubes and also feed in the immediate vicinity of either end of the open-ended tubes with their caudal extremities anchored within the tube. The 4 larval stages are followed by a black-colored pupal stage (lasting about 3 d) and emergence to a terrestrial adult (imago) stage. In the laboratory at 23 °C, and provided an adequate food supply, larvae will transition to a pupal stage in about 21 days, with males tending to pupate and emerge sooner than females (though generally overlapping). Adults do not feed, and typically live for 3 to 5 d (Sadler 1935) (159).

12.1.2.4 *Substrate Tolerance*—In the context of laboratory toxicity testing, larvae of *C. dilutus* appear to be tolerant of a wide range of substrates size and type. Several studies have shown that survival is not affected by particle size in natural sediments, sand substrates, or formulated sediments in both 10-d and long-term exposures (Ankley et al. 1994b (45); Suedel and Rodgers 1994 (35); Sibley et al. 1997b (167), 1998 (168)). Ankley et al. (1994b) (45) found that growth of *C. dilutus* larvae was weakly correlated with sediment grain size composition, but not organic carbon, in 10-d tests using 50 natural sediments from the Great Lakes. However, Sibley et al. (1997b) (167) found that the correlation between grain size and larval growth disappeared after accounting for inorganic material contained within larval guts and concluded that growth of *C. dilutus* was not related to grain size composition in either natural sediments or sand substrates. Emergence, reproduction (mean eggs/female), and hatch success were also not affected by the particle size composition of substrates in long-term tests with *C. dilutus* (Sibley et al. 1998) (168).

12.1.2.5 Organic Matter—Based on 10-d sediment tests, the content of organic matter in sediments does not appear to affect survival of *C. dilutus* larvae in natural and formulated sediments, but may be important with respect to larval growth. Ankley et al. (1994b) (45) found no relationship between sediment organic content and survival or growth in 10-d bioassays with *C. dilutus* in natural sediments. Suedel and Rodgers (1994) (35) observed reduced survival in 10-d tests with a formulated sediment when organic matter was <0.91 %; however, supplemental food was not supplied in this study, which may influence these results relative to the 10-d test procedures described in this standard. Lacey et al. (1999) (33) found that survival of *C. dilutus* larvae was generally not affected in 10-d tests by either the quality or quantity of synthetic (alpha-cellulose) or naturally derived (peat, maple leaves) organic material spiked into a formulated sediment, although a slight reduction in survival below the acceptability criterion (70 %) was observed in a natural sediment diluted with formulated sediment at an organic matter content of 6 %. In terms of larval growth, Lacey et al. (1999) (33) did not observe any systematic relationship between the level of organic material (for example, food quantity) and larval growth for each carbon source. Although a significant reduction in growth was observed at the highest concentration (10 %) of the leaf treatment in the food quantity study, significantly higher larval growth was observed in this treatment when the different carbon sources were compared at about equal concentrations (effect of food quality). In the latter study, the following gradient of larval growth was established in relation to the source of organic carbon: peat < natural sediment < alpha cellulose < leaves. Since all of the treatments received a supplemental source of food, these data suggest that both the quality and quantity of organic carbon in natural and formulated sediments may affect growth in sediment toxicity tests with *C. dilutus* (Lacey et al. 1999) (33). Vos et al. (2000) (39) evaluated several different types of food and found that a closely related species, *C. riparius*, grew fastest and largest on animal-based foods (including the tropical fish flake food used in sediment toxicity tests described here) compared to plant-based foods, differences attributed to differences in chemical composition.

12.2 General Culturing Procedures:

12.2.1 Performance Based Evaluation—Acceptability of a culturing procedure is based largely on performance of organisms in culture and in the sediment test. No single technique for culturing test organisms is required. What may work well for one laboratory may not work as well for another laboratory. While 12.3 and 12.4 describe a variety of culturing procedures that have been successfully used, the test of culture adequacy is to produce organisms in sufficient numbers and in health sufficient to meet the test acceptability criteria listed in Table A1.1, Table A2.1, Table A3.1, and Table A4.1.

12.2.2 Source and Verification—All organisms in a test must be from the same culture cohort. Organisms may be obtained from laboratory cultures or from an outside source, provided the outside source can provide documentation that the cohort of organisms meets applicable requirements. The species of test organisms should be positively identified using an appropriate

taxonomic key, and verification should be documented (Pennak 1989 (135), Merritt and Cummins 1996) (136). Obtaining organisms from wild populations should be avoided unless organisms are cultured through several generations in the laboratory. The ability of the wild population to cross breed with the existing laboratory population should be determined (Duan et al. 1997) (137). The sensitivity of the wild population to select reference chemicals should also be documented.

12.2.3 Organisms from Outside Cultures—To the extent practicable, in-house culture of test organisms is recommended. In-house culture provides a laboratory greater control and understanding of culture health, and avoids the risks of stressing organisms in shipment. However, use of test organisms obtained from commercial sources is allowed. Organisms from outside cultures should be shipped in well-oxygenated water in insulated containers to maintain temperature during shipment. Temperature and dissolved oxygen of the water in the shipping containers should be measured on arrival to determine if the organisms might have been subjected to low dissolved oxygen or temperature fluctuations. The temperature of the shipped water should be gradually adjusted to the desired culture temperature at a rate not exceeding 1 °C every 1 to 2 h. Additional reference-toxicity testing is suggested if organisms are not cultured at the testing laboratory (11.10).

12.2.3.1 Verification of Age—Because this third edition tightens the age/size requirements for organisms used to start sediment toxicity tests, it is important that organisms obtained from outside sources have good documentation of their history and age. For *C. dilutus*, the 10-d test method (Annex A2) indicates that organisms should not be >0.12 mg ash-free dry weight at the start of the study. While it may be impractical to cancel a test if larvae arrive from an outside source and exceed that size specification, laboratories receiving midge larvae larger than 0.12 mg AFDW should work with their supplier to ensure that future organisms are younger, and can meet the ≤0.12 mg AFDW specification. For midge used for life-cycle testing starting with 3-d old larvae (Annex A4), it is desirable to ship the organisms as egg masses so that hatching, health, and age can be more readily tracked by the testing laboratory.

12.2.3.2 Sieving *H. azteca* from Mixed Age Cohorts No Longer Acceptable—Previous versions of this standard have allowed sieving mixed age cultures as a means to isolate a cohort of *H. azteca* for use in toxicity testing. This option is no longer included, because experience has shown sieving is not as selective for age as originally thought. Further, the refinements of the *H. azteca* test methods require a tighter age window to minimize variability and ensure synchronous reproduction in the life-cycle test (Annex A3). Laboratories obtaining test organisms from outside sources should verify with their supplier that the cohort of amphipods was not isolated by sieving.

12.2.4 Monitoring Cohorts of Test Organisms—The cohort of organisms used for any test should appear healthy, behave normally, feed well, and have low mortality in cultures, and during pre-test holding (for example, <10 % for 48 h before the start of a test). If the organisms fail to meet these criteria, the entire cohort should be discarded and a new cohort should be obtained. All organisms in the cohort should be monitored;

high variation in size among a known age cohort may be a sign of culture problems. Test organisms should be handled as little as possible. When handling is necessary, it should be done as gently, carefully, and as quickly as possible.

12.2.5 Culture Water—*Chironomus dilutus* can be cultured in a variety of waters. Water of a quality sufficient to culture fathead minnows (*Pimephales promelas*) or cladocerans will generally be adequate. However, *H. azteca* of the U.S. laboratory clade have additional requirements of ≥ 0.04 mg Br/L and ≥ 15 mg Cl/L (7.1). Water supplies that are otherwise suitable but have less Br or Cl than required can be amended by adding NaBr and NaCl to achieve sufficient concentrations.

12.2.6 Temperature and Light—Unless there is a study-specific need for another temperature, cultures should be maintained at the same average temperature as the test methods, 23 °C. Daily fluctuations of more than 2 °C should be avoided. There is no indication that *H. azteca* or *C. dilutus* are sensitive to the light intensity in culture; an illuminance between 100 to 1000 lux is common.

12.2.7 Physical and Chemical Monitoring of Culture Water—The appropriate intensity of chemical/physical monitoring of water quality in cultures depends on the nature of the water source and culture systems and their susceptibility to fluctuations, the length of time that cultures have been maintained under consistent methods, and the importance of the parameters. Temperature should be measured daily; it is recommended that cultures incorporate real-time monitoring equipment that can alert culture staff to temperature excursions if they occur. Maintenance of dissolved oxygen is important to culture health, and warrants careful consideration, particularly for midge, as the ration provided to larval growth tanks often creates heavy oxygen demand; weekly monitoring is appropriate, but more frequent measurements should be made when procedures are changed, or if water turbidity increases (indicating microbial growth) or unusual amounts of biological growth are observed on surfaces. Ammonia is a byproduct of protein metabolism, and may be elevated in cultures fed animal-based foods (for example, tropical fish flakes or trout chow). Ammonia should be monitored frequently when cultures are initially developed, until there is confidence that ammonia buildup is not a problem. Additional ammonia monitoring should also be triggered by observation of excessive microbial growth. Unless there is specific reason to believe there may be unexpected variation, measures of pH, hardness, alkalinity, and conductivity can be relatively infrequent (for example, quarterly). If organisms are cultured in reconstituted water or other waters that are prepared in batches, chemical characterization of each batch should be conducted to ensure that the batch was prepared in a way consistent with previous batches.

12.2.8 Monitoring for Contaminants—Insuring that water, substrates, and food are free of undesirable contaminants is important, but the appropriate parameters and frequency are dependent on the types and stability of the sources for these materials. Having at least initial characterization of background contamination is desirable. Lists of suggested analytes are provided by Guide E729.

12.2.9 Culture Performance Monitoring—Records should be kept to verify culture performance and provide a basis for identifying and troubleshooting changes in performance. Examples of culture monitoring practices are provided for each species in the example culture performance described within 12.3 and 12.4.

12.2.10 Reference Toxicant Testing—It may be desirable for laboratories to periodically perform 96-h water-only reference-toxicity tests to assess the sensitivity of culture organisms. A more detailed discussion of reference toxicant testing is provided in 11.10.

12.3 Culturing Procedures for *Hyalella azteca*:

12.3.1 General—This standard describes 2 test methods using *H. azteca*, a 10-d exposure beginning with 7 to 10-d old amphipods (Annex A1) and a 42-d life cycle test beginning with 7 to 8-d old amphipods (Annex A3). Both tests require organisms of known age, and the culture methods to produce organisms for either test are largely the same. Cohorts for either test are generally obtained by isolating adult organisms of reproductive age, then collecting the young produced over the subsequent 24 h. These young are isolated from the adult organisms and reared as a cohort of known age organisms. The adult cultures may consist of individuals of a known age, or a mixed age culture from which adults are removed as needed to obtain young.

12.3.2 Sieving from Mixed Age Cohorts No Longer Acceptable for Test Cohorts—Previous versions of this standard have allowed sieving smaller organisms from mixed age cultures as a means to isolate a cohort of *H. azteca* for use in toxicity testing. This option is no longer included, because experience has shown sieving is not as selective for age as originally thought. Further, the refinements of the *H. azteca* test methods require a tighter age window to minimize variability and ensure synchronous reproduction in the life-cycle test (Annex A3). While sieving is not an acceptable means to obtain a cohort for testing, sieving may be used to isolate adult organisms from a mixed age culture, with the isolated adults then being held for capture of young produced.

12.3.3 Example Methods for Culturing *H. azteca*—As indicated above, laboratories may use a variety of culture methods for *H. azteca*, as long as they can be shown to sustain an ongoing culture population that is able to produce organisms of sufficient numbers and in health sufficient to meet the test acceptability criteria. 12.3.3.1 to 12.3.3.6 describe methods that have been used successfully in particular laboratories; these are not provided as strict specifications, but instead as examples that may aid laboratories in designing their own culture methods.

12.3.3.1 *Hyalella azteca* culture at the U.S. EPA laboratory in Duluth, MN:

(1) Culture Water—Two waters have been used to culture *H. azteca*, both modifications of sand-filtered and UV-treated Lake Superior water to which salts are added to modify the composition. One water has a hardness and alkalinity of about 110 and 70 mg/L as CaCO₃ (respectively), with nominal ion concentrations of 14.8 mg Na/L, 2.7 mg K/L, 30.5 mg Ca/L, 9.0 mg Mg/L 19.4 mg Cl/L, and 45.2 mg SO₄/L. The second water has a hardness and alkalinity of about 110 and 70 mg/L as CaCO₃ (respectively), with nominal ion concentrations of 14.8 mg Na/L, 2.7 mg K/L, 30.5 mg Ca/L, 9.0 mg Mg/L 19.4 mg Cl/L, and 45.2 mg SO₄/L.

CaCO₃, and a nominal ion composition of 11.2 mg Na/L, 1.5 mg K/L, 14.6 mg Ca/L, 4.1 mg Mg/L, 15.0 mg Cl/L, and 14.9 mg SO₄/L. NaBr is added to achieve a nominal Br concentration of 0.04 mg/L.

(2) *Foods*—Rinsed *Conticribra weissflogii* (formerly known as *Thalassiosira weissflogii*; 12.5.2.1) suspension (6 g/L solids) and TetraMin® flakes crushed dry to pass through a #50 US standard sieve and suspended in water at 1g/100 mL.

(3) *Substrate*—Cotton gauze (Johnson and Johnson sterile 4 × 4 in. gauze sponges). Gauze sponges are unfolded to produce an overall dimension much greater than 10 cm by 10 cm.

(4) *Adult Culture*—Adult *H. azteca* are cultured in 20-L glass aquaria containing 8 L of culture water, 3 gauze sponges (unfolded), and stocked with about 500 adult amphipods. Water is aerated gently, and changed once/week. Each tank is provided 10 mL of diatom solution and 7 mL of Tetramin flake solution at water renewal and daily thereafter.

(5) *Obtaining Known-age Cohorts*—Adult *H. azteca* are isolated from the adult culture tanks by first agitating the gauze substrate to dislodge the adults, then pouring the contents of the tank through a #25 US standard size set into a pan so that it does not go dry. The culture tank is rinsed with culture water to dislodge any remaining amphipods, then the culture tank is scrubbed and refilled with culture water, fresh substrate, and food. The adults retained on the sieve are then returned to the culture tank. The following day, the process is repeated, but the young that pass the sieve are retained as a cohort. Young from multiple adult culture tanks are combined, then apportioned to rearing vessels.

(6) *Rearing Known-age Cohorts*—Young amphipods are reared in 2-L glass battery jars containing 1 L of culture water, 1 gauze sponge (unfolded), and 1.5 mL of diatom solution and 1 mL of Tetramin flake solution daily. Jars are aerated gently. Once per week, 1 jar containing 150 amphipods 8 to 9-d old is established to rear organisms to reproductive maturity to supplement adult cultures. These jars are renewed weekly in a manner similar to that described for adult cultures.

(7) *Culture Performance Monitoring*—Several measures of culture performance are made:

(a) Each week, hand counts of adults are performed in 2 adult culture tanks to monitor adult survival.

(b) Each week, hand counts of young production are made in 2 adult culture tanks.

(c) Each week, 1 battery jar containing 250 neonates is monitored for young survival (counted again after 8 days).

(d) Each week, 1 battery jar containing 30 adult male and 30 adult females is established to track adult survival and young production over a 2 week period.

(e) Reference toxicant tests are conducted quarterly.

12.3.3.2 *Hyaella azteca* culture at the US Geological Survey in Columbia, MO:

(1) *Culture Water*—Well water with a hardness of about 280 mg/L as CaCO₃ and alkalinity of about 250 mg/L as CaCO₃, pH about 8.2, and with ion concentrations of about 24 mg Na/L, 2.4 mg K/L, 70 mg Ca/L, 27 mg Mg/L, 22 mg Cl/L,

60 mg SO₄/L, and 0.09 mg Br/L. Well water diluted with deionized water to a hardness of 100 mg/L as CaCO₃ is often used to grow out young amphipods for testing.

(2) *Foods*—The mass cultures of amphipods are fed an amphipod Flake food mix and maple leaves (soaked in culture water for 2 weeks before use). The amphipod flake food mix components include: (1) 800 g crushed flake food (Worldwide Aquatics); (2) 100 g Otohime B1/B2 (Reed Mariculture); and (3) 100 g MDA (midge dried algae). The MDA components include: (1) 5 g alfalfa meal (Bourne Feed); (2) 4 g Alga Mac 3050 Flake + DHA (AquaFauna BioMarine); (3) 1 g AlgaMac Enhance (AquaFauna Biomarine); (4) 16.5 g each of dry *Spirulina* and *Chlorella* (Bulk Foods); (5) 2 g Beta Meal (NE Brine Shrimp); and (6) 5 g Kelp Meal (Florida Aquafarms). All of the ingredients in the amphipod flake food mix are mixed dry in a container and stored at room temperature. Before use, the dry Flake Food Mix and the Midge Dried Algae are wetted with enough well water to create a slurry and is then held for use at about 4 °C for up to 3 weeks.

(3) *Substrate*—Maple leaves (about 20 cm in diameter; soaked in well water for about a month). Leaf replacement every 2 to 3 weeks.

(4) *Adult Culture*—Adult *H. azteca* are cultured in 38-L glass aquaria containing 34 L of culture water, about 10 maple leaves, fiberglass window screen (3 to 4 pieces of 10 × 10 cm wire gray fiberglass screen cloth; and stocked with about 500 adult amphipods. Water is aerated gently and each aquarium receives about 10 volume additions/day of well water. Each aquarium is provided 6 g daily of amphipod flake food mix.

(5) *Obtaining Known-age Cohorts*—Using a soft mesh fish net, gently net large amounts of mixed-age amphipods into a No. 25 (710-µm mesh) sieve that has been placed in a round Nalgene bowl with culture water close to top of sieve thus allowing enough volume of water for amphipods to swim. Gently aerate and allow to sit for 15 minutes. Immature amphipods will pass through the sieve. After 15 minutes, transfer the sieved adult amphipods to a second Nalgene bowl filled with culture water. Gently aerate and add a few small leaves. After 24 h, lift sieve with adults up and down several times to rinse <24 h-old amphipods through sieve. Remove sieve and place adults back into culture chambers. Transfer water and <24-h old amphipods from the Nalgene bowl into a 6-L round plastic bucket. Fill to 4 L with test water. Within each bucket place 1 to 2 leaves, set in temperature controlled water bath with aeration and flowing test water (about 7 volume additions/day). The number of newborn amphipods is estimated to determine if adequate numbers have been collected for the test. Isolation of about 4000 mixed-age adults (some in amplexus and others not in amplexus) provides about 800 newborn amphipods in 24 h and required less than 1 man-hour of time.

(6) *Rearing Known-age Cohorts*—About 1500 to 2000 young amphipods before the start of a test are reared in 6-L plastic buckets containing 4 L of test water and 2 maple leaves. The water in the buckets is aerated gently with about 7 volume additions/day of test water using a continuous slow drip from a valve. The buckets have a 2.5 × 10 cm opening covered with 150-µm polypropylene mesh to allow for water circulation.

water flow during the grow-out of the young amphipods. Each bucket is provided 36 mg dry weight diatoms each day, plus an increasing ration of Tetramin slurry (10 mg dry weight days 0-3; 20 mg dry weight on day 4; 40 mg dw on days 5 to 7).

(7) *Culture Performance Monitoring*—Feeding records and temperature recorded daily. Aeration checked daily. Reference toxicant tests are conducted periodically.

12.3.3.3 *Hyalella azteca* Culture at the Illinois Natural History Survey Laboratory in Champaign, IL:

(1) *Culture Water*—Two waters have been used to culture *H. azteca*—one with a hardness of about 110 mg/L as CaCO₃, prepared according to a formula developed by Borgmann (1996 (46); 7.1.3.5(1)), and another with a nominal hardness of 100 mg/L as CaCO₃, prepared according to a formula developed at the USEPA laboratory in Duluth, MN, (hereafter referred to as Duluth 100; 7.1.3.5(3)). Both are reconstituted waters made by adding salts to deionized water. The “Borgmann” water has nominal ion concentrations of 23 mg Na/L, 2.0 mg K/L, 40 mg Ca/L, 6.1 mg Mg/L, 73 mg Cl/L, 24 mg SO₄/L, and 0.8 mg Br/L. The Duluth 100 water has a nominal ion composition of 34 mg Na/L, 3.9 mg K/L, 285 mg Ca/L, 7.7 mg Mg/L, 28 mg Cl/L, 59 mg SO₄/L, and 0.04 mg Br/L.

(2) *Foods*—Rinsed *Conticribra weissflogii* (formerly known as *Thalassiosira weissflogii*; 12.5.2.1) suspension (8 to 10 mg (dry weight)/mL) and TetraMin® flakes crushed dry to pass through a #35 US standard sieve (<500 µm) or a #60 sieve (<250 µm) and suspended in deionized water at 1 mg/1 mL.

(3) *Substrate*—Sterile cotton gauze (Dynarex sterile 3 × 3 in. gauze sponges). Gauze pads are unfolded to produce an overall dimension much greater than 7.5 cm by 7.5 cm.

(4) *Adult Culture*—Adult *H. azteca* (about 500/container) are cultured in 9.5-L glass aquaria containing about 4 L of culture water, and several gauze sponges (unfolded). Water is aerated gently, and changed from multiple times weekly (if collecting known age young) to about every 2 weeks (for general maintenance). Each tank is provided about 50 mg of <500 µm Tetramin flake 3 times/week. Cotton gauze pads are replaced as needed.

(5) *Obtaining Known-age Cohorts*—The US Standard sieves are nested with a #18 on top, #40 in the middle, and a #60 on the bottom. The entire contents of an adult culture tank are poured into the top sieve. The culture tank is refilled with culture water. Any gauze pieces are thoroughly rinsed into the top sieve and then discarded, to be replaced by fresh gauze. Any smaller *H. azteca* are rinsed through the top sieve using water. Larger individuals remaining in the top sieve are rinsed back into the culture tank. This process is then repeated with the middle (#40) sieve, rinsing the smallest individuals through to the bottom sieve. The contents of the middle sieve are returned to the culture tank and the individuals in the bottom (#60) sieve are discarded. Two days later, this entire process is repeated but the contents of the #60 sieve are rinsed into a 1-L beaker containing culture water.

(6) *Rearing Known-age Cohorts*—Young amphipods are reared in 1-L glass beakers containing 1 L of culture water, 1 gauze sponge (unfolded), and fed 5 mg (as dry weight but

added as solution) of diatom solution and 3 mg of <250 µm Tetramin flake 3 times/week. Beakers are aerated gently. Water in these beakers is renewed weekly as needed.

(7) *Culture Performance Monitoring*—Reference toxicant tests are conducted as needed.

12.3.3.4 *Hyalella azteca* Culture at the US Army ERDC (USACE) in Vicksburg MS:

(1) *Culture Water*—Carbon filtered dechlorinated municipal water (Vicksburg, MS) with a hardness of about 70 mg/L as CaCO₃ and alkalinity of about 85 mg/L as CaCO₃, pH about 8.4, and with ion concentrations of about 20 mg Na/L, 2.1 mg K/L, 11 mg Ca/L, 10 mg Mg/L, 18 mg Cl/L, 8 mg SO₄/L, and <1 mg Br/L.

(2) *Foods*—Mass cultures of amphipods are fed Tetramin flake fish food which is ground to a powder and placed in a water suspension. Cultures are also provided maple leaves as substrate.

(3) *Substrate*—Dried maple leaves are conditioned by soaking in saltwater and freshwater before use. Leaves are replaced every 2 to 3 weeks. Nitex plastic mesh is also provided as an inert substrate in mass cultures.

(4) *Adult Culture*—Adult *H. azteca* are cultured in 38-L aquaria containing 34 L of culture water, enough maple leaves to cover the bottom of the aquaria, and Nitex plastic mesh of various sizes and shapes. Aquaria are initially stocked with about 500 adult amphipods. Water is aerated gently and each aquarium receives about 12 volume additions/d of carbon filtered dechlorinated municipal water. 500 mg flake food is provided 3 times weekly.

(5) *Obtaining Known-age Cohorts*—Using a mesh fish net, a large number of mixed-age amphipods are gently netted from the adult mass culture and rinsed on a 30-cm No. 25 (710-µm mesh) sieve that has been placed in a polypropylene tray (45 cm by 33 cm by 13 cm) with culture water close to top of sieve thus allowing enough volume of water for amphipods to swim. Maple leaves and other debris are separated from the adult amphipods. After allowing 15 minutes of settling, the sieve is moved gently up and down in the water column to encourage immature amphipods to pass through the sieve. The sieve containing the adult amphipods is transferred to a second polypropylene tray filled with culture water. Gentle aeration and a few maple leaves are added to the sieve. After 24 h, the sieve is lifted up and down several times to rinse <24 h-old amphipods through sieve. The adults are transferred back to the mass culture, and the <24-h old amphipods from the polypropylene tray are transferred into 4-L glass jars with 3.5 L culture water. Five to 10 maple leaves are added to each jar, which is held in a temperature-controlled water bath or environmental chamber with aeration.

(6) *Rearing Known-age Cohorts*—About 1000 young amphipods are reared in 4-L glass jars containing 3.5 L of culture water and 5 to 10 maple leaves and are fed 30 mg flake food daily. If food build up is observed feeding frequency is decreased. The water in the jars is aerated gently and a 50 % water exchange is conducted 3 times weekly.

(7) *Culture Performance Monitoring*—Water quality parameters (temperature, pH and dissolved oxygen) are measured weekly. Feeding and daily observation logs are maintained. Reference toxicity testing (96-h survival test using KCl) are conducted and control charts are updated monthly.

12.3.3.5 *Hyaella azteca* Culture at Smithers Viscient, LLC, Wareham MA:

(1) *Culture Water*—Well water with a hardness of 60 to 80 mg/L as CaCO₃ and alkalinity of about 20 mg/L as CaCO₃, pH 7.0, and conductivity of 400 to 600 µS/cm.

(2) *Foods*—The main culture is fed dry flaked fish food (Sera Vipan Staple), daily *ad libitum*. Known-age offspring which are isolated for testing are initially fed a combination of flaked fish food slurry (100 mg/mL prepared in laboratory well water), *Ankistrodesmus falcatus* (unicellular green algae), and YCT. YCT is then provided daily during rearing.

(3) *Substrate*—Nylon screen of varying mesh sizes (for example, 10 to 40 US mesh).

(4) *Main Culture*—Mixed-age *H. azteca* are cultured in 38-L glass aquaria containing 31 L of culture water, at a density of about 1000 amphipods/aquarium. Several layers of nylon screen, cut into pieces about 100 to 150 cm² are placed in each aquarium as substrate. Each aquarium is maintained under flow-through conditions at a renewal rate of 5 to 10 volume additions/d and the water within each aquarium is continuously aerated. Dry flaked fish food is provided daily to each aquarium *ad libitum*, typically enough to lightly cover the surface of the water.

(5) *Obtaining Known-age Cohorts*—Amphipods from the main culture are removed and rinsed over an 18-mesh sieve (1-mm opening). Adult amphipods retained on the sieve are then transferred to 9.5-L aquaria containing about 8 L of well water and additional nylon substrate. Several hundred adults are stocked in each aquaria and maintained under static conditions with aeration for the following 24 h. Food is not provided during this holding period. Offspring are collected on the day following adult isolation to provide <24 h-old neonates.

(6) *Rearing Known-age Cohorts*—On the day following adult isolation, offspring are pipetted from the aquaria and transferred to 1-L glass beakers containing 0.8 L of well water. Each beaker is stocked at a rate of 200 to 250 offspring and is maintained under static conditions with aeration. A single piece of nylon substrate (20 to 25 cm²) is also provided. On the first day of isolation, about 100 mg of slurried flake fish and a small amount of *Ankistrodesmus falcatus*, a unicellular green algae, are added to each beaker. In addition, 1.0 mL of YCT is provided daily during holding.

(7) *Culture Performance Monitoring*—Feeding records and temperature are recorded daily. Dissolved oxygen and pH are measured weekly in the main culture, and every other day during the rearing of known-age amphipods for testing. Reference toxicant tests are conducted periodically with potassium chloride.

12.3.3.6 *Hyaella azteca* Culture at the EAG Laboratories-Easton MD:

(1) *Culture Water*—Water used for holding the test organisms are obtained from a well located on the EAG Laboratories site in Easton, Maryland. The water is sand filtered, treated to remove any particulate matter over 0.45 µm particle size and UV sterilized before use. The water typically has hardness of about 140 mg/L and alkalinity 180 mg/L as CaCO₃, specific Conductance of about 365 µS/cm, and total organic carbon of less than 2 mg/L. These measurements are taken weekly to monitor the consistency of the well water. In the culture unit, the water is supplemented with sodium chloride (NaCl) and sodium bromide (NaBr) by delivering a solution into a reservoir where it is mixed with well water to provide an addition of 0.02 mg/L Br, 20 mg/L Cl, and then into each culture aquaria through calibrated flow.

(2) *Foods*—Ground Tetramin® flake food fed about 0.5 g 3 times/week.

(3) *Substrate*—2 to 3 pieces of Cotton surgical gauze unfolded. The gauze is replaced with new gauze twice a week.

(4) *Adult Culture*—Culture consists of mixed age *H. azteca* in a 20-L glass aquaria containing about 16 L of well water under flow-through conditions. The flows to each tank are calibrated/verified at least weekly and are adjusted as needed.

(5) *Obtaining Known-age Cohorts*—Adult *H. azteca* are isolated from the mixed age culture aquaria by first agitating the gauze substrate to dislodge the organisms, then pouring the contents of the aquaria through a sieve into a pan. Appropriate sieve sizes for collecting adult *Hyaella* are 500, 600 or 710 µm. The organisms that are retained by this sieve are rinsed with well water and placed into plastic tub containing gently aerated culture water overnight. The following day, the organisms on the sieve will be returned to the culture aquaria and the young that pass through the sieve are retained and are considered to be <24 h old. Young from multiple collection tubs are sieved using a 125 µm sieve (to obtain young of a consistent size) and are carefully rinsed into an appropriate holding container.

(6) *Rearing Known-age Cohorts*—Less than 24-h old neonates are reared in 2-L glass beakers containing 1.8 L of culture water until appropriate age to be used for testing (7-8 d old). These beakers are held in 20-L culture aquaria containing culture water which serves as a water bath to maintain the temperature. One piece of surgical gauze is added to the beaker. The beakers are gently aerated.

12.4 Culturing Procedures for *Chironomus dilutus*:

12.4.1 *Cohort Requirements*—This standard describes 2 test methods using *C. dilutus*, a 10-d exposure beginning with 7 to 10-d post hatch larvae ([Annex A2](#)) and a life cycle test beginning with 3-d old midges ([Annex A4](#)). The organism specifications for both tests have changed from the previous edition of this standard. For the 10-d test, under the previous specification based on age and instar, some laboratories were observing pupation during the 10-d test. It is important that larvae in a 10-d test remain in a larval growth phase and do not reach pupation during the test. For this reason, a size specification of ≤0.12 mg ash-free dry weight (AFDW)/individual (on average) is now used, which is generally 2nd instar. If cohorts of organisms for 10-d tests are larger than 0.12 mg AFDW,

even occasionally, then the culture procedure should be adjusted to provide cohorts at a younger age (and consequently smaller size). For the life-cycle test ([Annex A4](#)), the revised test method described in this version of the standard begins with 3-d old larvae, rather than the <24-h old larvae in the previous version of the standard (E1706-19) and in USEPA (2000) (8). Both tests require cohorts within 24-h in age that are produced from a group of egg masses.

12.4.1.1 Definition of Age for *C. dilutus*—Defining the age of larval cohorts obtained in this way is complicated by 2 factors: (1) hatched larvae remain in the egg mass for a period (typically 12 to 24 h) before leaving the egg mass and becoming free-living; and (2) in culture vessels seeded with egg masses, the exact time of hatch (or leaving the egg mass) may not be readily observable. For this reason, age of larval cohorts is indexed to the time they would be expected to leave the egg mass, based on normal development timelines (typically 72 to 96 h after eggs are laid).

12.4.2 General—Because *C. dilutus* has both aquatic and terrestrial (aerial) life stages, culture is organized similarly. Larvae are generally reared in large groups as a same-age cohort, progressing through 4 instars to an aquatic pupal stage in about 20 d, and then emerge as flying adults from the water surface. Cultures are generally covered with some form of netting to prevent escape of the emerged adults. Adult females and males are captured and held together in a flask or other container with water at the bottom. If mates are available, adult midges usually copulate within 24 h of emerging, then females deposit an egg mass on the water, which quickly hydrates into a gelatinous mass with eggs suspended within in a coiled arrangement. The adults typically live 3 to 5 d and do not feed. Egg masses have a fair structural integrity and can be easily transferred and manipulated. The next generation is generally started by combining partial egg masses from several females to increase genetic diversity. At 23 °C, eggs hatch in about 3 d, then the larvae remain in and feed on the egg mass for 12 to 24 h before leaving.

12.4.3 Example Methods for Culturing *C. dilutus*—As indicated above, laboratories may use a variety of culture methods for *C. dilutus*, as long as the methods can be shown to sustain an ongoing culture population that is able to produce organisms of sufficient numbers and in health sufficient to meet the test acceptability criteria. Sections 12.4.3.1 to 12.4.3.5 describe methods that have been used successfully in particular laboratories; these are not provided as strict specifications, but instead as examples that may aid laboratories in designing their own culture methods.

12.4.3.1 *Chironomus dilutus* Culture at the U.S. EPA laboratory in Duluth, MN:

(1) **Culture Water**—Midges are cultured in sand-filtered and UV-treated Lake Superior water; this water has a hardness and alkalinity of about 45 and 43 mg/L as CaCO₃ (respectively), with nominal ion concentrations of 0.6 mg Na/L, 2.7 mg K/L, 14.0 mg Ca/L, 2.9 mg Mg/L, 1.5 mg Cl/L, and 3.4 mg SO₄/L.

(2) **Food**—A 56 g/L solution of Tetramin® Tropical Fish Flakes, blended into a fine suspension.

(3) **Substrate**—A layer of clean sand, about 425 µm grain size in an even layer about 5 mm deep.

(4) **Larval Rearing**—Cultures are started in 20-L glass aquaria, fitted with an overflow standpipe to achieve a volume of 6 to 8 L of culture water. A layer of sand is added to the bottom, and each tank has an airstone providing moderate aeration (the midge food exerts a substantial oxygen demand, and maintaining dissolved oxygen is important). Each tank is seeded with partial egg masses from multiple females (for example, 1/3 of an egg mass from each of 3 females) and each tank is fed daily with 5 mL of the Tetramin solution. A water supply on an automated timer provides 1 volume addition/d of fresh water over the course of about 2 h. Addition of fresh water is important to the maintenance of acceptable water quality in the cultures. Tanks prepared in this way typically yield about 500 larvae.

(5) **Collection of Adults, Mating, and Egg Collection**—When larvae begin to pupate, the culture tank is covered with a piece of nylon mesh (panty hose) to keep emerged adults confined. A collection and breeding flask ([Fig. 4](#)) is used to capture adults from the culture and allow them to breed. These are 1-L Ehrlenmeyer flasks containing a rectangular piece of coarse (for example, window screen size) nylon screen. A two-hole stopper is fitted with 2 glass tubes connected to plastic tubes, 1 for collecting the adults (by aspiration) and the other providing suction (by mouth). Thirty females and 10 males are collected by aspiration from the culture tanks, then about 100-mL of culture water is added to the bottom of the flask. Male midges tend to be more slender and have plumous (feather-like) antennae, while females have larger abdomens and small slender antennae. After 24 h, adults have usually mated and will have deposited egg masses in the water, though egg deposition may continue for as much as 72 h. Egg masses can be collected and transferred using glass tubes of appropriate diameter, and can be incubated in small containers (beakers, crystallizing dishes, or Petri dishes) if needed.

(6) **Rearing Known-age Cohorts**—Cohorts of larvae for 10-d tests ([Annex A2](#)) are reared just as described in 12.4.3.1(4); tests are started with 8-d old larvae, which routinely meet the ≤0.12 mg AFDW standard. The procedure for obtaining 3-d old larvae for life-cycle testing ([Annex A4](#)) is described in [A4.2.7.4](#) and [A4.2.7.5](#). Briefly, egg masses are held individually in small dishes until larvae begin to emerge from the egg mass, then are reared on a solution of algae (dried *Chlorella*, reconstituted in water at 10 g/L) fed at a rate of 1.25 mL daily to dishes containing 200 mL culture water. The algal particles provide both food and detritus for case construction.

(7) **Culture Performance Monitoring**—Several measures of culture performance are made:

(a) Wet chemistry (temperature, pH, ammonia, and DO) are monitored weekly, or more often if observations suggest there may be problems (for example, excessive microbial growth). Temperature of the culture water supply and room air temperature are also continuously monitored.

(b) Egg mass viability and approximate hatching success is observed for every egg collection; in addition, percent hatch is quantitatively measured on 6 to 9 egg masses per cohort.

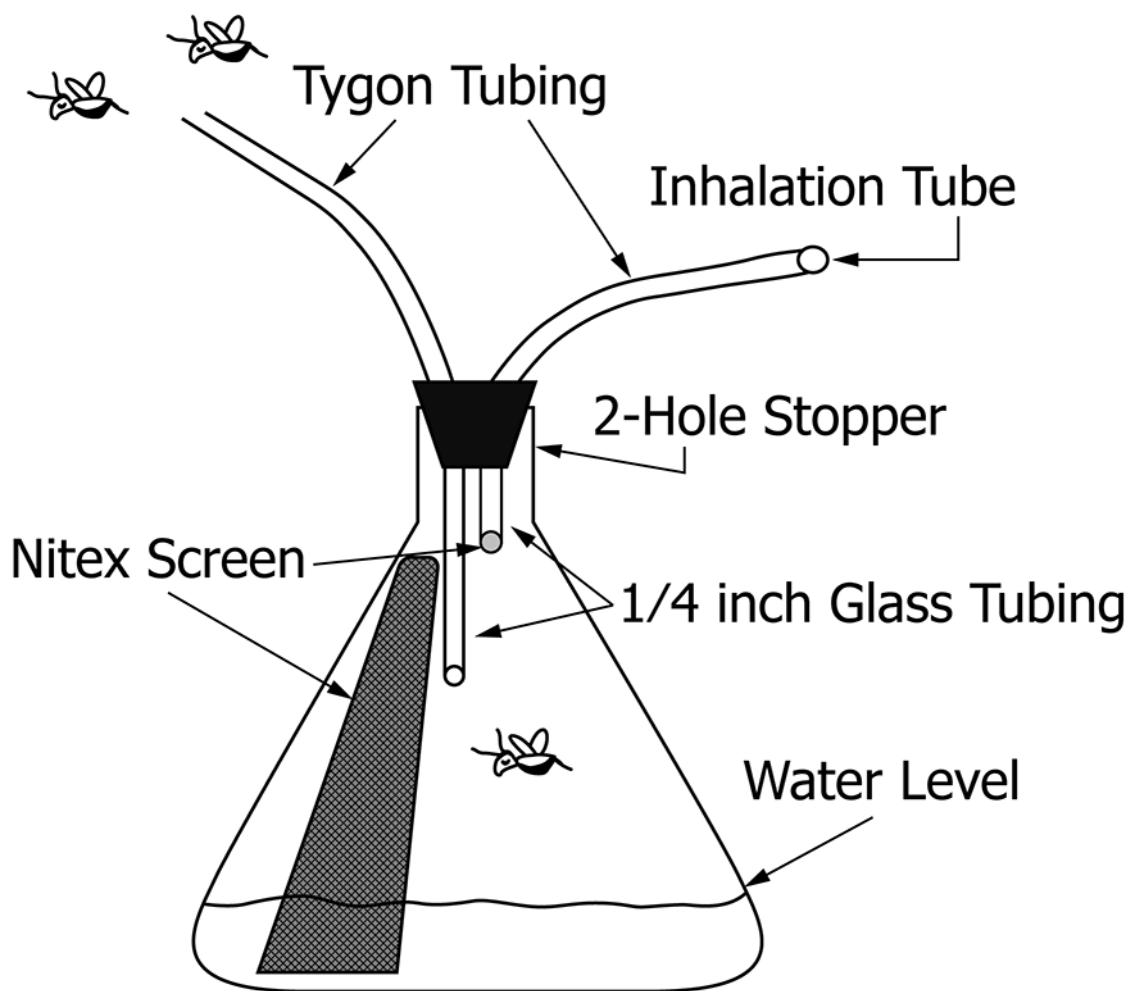


FIG. 4 Diagram of Adult Collection and Breeding Flask Used in Culture of *C. dilutus*

(c) Each month, subsamples of 8-d old larvae are collected for weight and length determination. Twenty larvae are evaluated for length and instar (based on head capsule width), and 40 are used for AFDW determination.

(d) On Mondays and Fridays, adult emergence and sex ratio is determined on a subset of culture tanks. The prior day, all adults are removed from the headspace above the culture tank, then the next day the number and sex of all emerged adults is determined and recorded.

(e) Reference toxicant tests are conducted quarterly.

12.4.3.2 *Chironomus dilutus* culture at the U.S. Geological Survey in Columbia, MO:

(1) *Culture Water*—Well water with a hardness of about 280 mg/L as CaCO_3 and alkalinity of about 250 mg/L as CaCO_3 , pH about 8.2, and with ion concentrations of about 24 mg Na/L, 2.4 mg K/L, 70 mg Ca/L, 27 mg Mg/L, 22 mg Cl/L, 60 mg SO_4 /L, and 0.09 mg Br/L. Well water diluted with deionized water to a hardness of 100 mg/L as CaCO_3 is often used to grow out young midges for testing.

(2) *Foods*—The mass cultures of midges are fed a combination of 3 food mixtures: (1) midge algal food mix (MAFM); (2) midge dried algae mix (MDAM)—and (3) midge flake food

mix (MFFM). The MAFM is comprised of: (1) 10 mL Shellfish Diet (Reed Mariculture); (2) 5 mL Nanochloropsis (Reed Mariculture); (3) 15 mL Selenastrum (Aquatic Biosystems); (4) 5 mL Phyto-Feast (Reed Mariculture); and (5) 20 mL of well water. The MDAM is comprised of: (1) 5g Alfalfa Meal (Bourne Feed); (2) 4 g Alga Mac 3050 Flake + DHA (AquaFauna BioMarine); (3) 1 g AlgaMac Enhance (AquaFauna BioMarine); (4) 6.5 g each of *Spirulina* and *Chlorella* (Bulk Foods); (5) 2 g Beta Meal (NE Brine Shrimp); and (6) 5g Kelp Meal (Florida Aquafarms) mixed with well water to create a thin paste. The MFFM is comprised of: (1) 125 g MDA; (2) 300 g crushed flakes (Worldwide Aquatics); and (3) 250 g Alfalfa Meal (Bourne Feed) mixed with well water to create a medium to thick slurry. Cohorts of known age midges for starting toxicity tests are reared on a combination of crushed (<250 μm) Tetra-min flaked fish food in a 10 mg/mL water suspension, and a 1.25 mg/mL suspension of *Chlorella* (alga) powder (Bulk Food).

(3) *Substrate*—None (added food provides substrate for case building).

(4) *Larval Rearing*—Mass cultures of midges are housed in 38-L glass aquaria with about 10 L of water. The

aquarium has a plexiglass lid modified by cutting a semicircle into the lid 17.75 cm across by 12.5 cm. Clear polypropylene mesh (20 mesh/0.4 cm) is cut to size and melted to the plastic lid. The screen provides air exchange and retains emerging adults. Two holes about 0.6 cm in diameter are drilled through the uncut portion of the lid to provide access for an airline and water line. Midges are cultured under flow through conditions with moderate aeration. Each container is seeded with about 2 egg masses each day from multiple females that were collected the previous day. A total of 7 culture aquaria are maintained. One aquarium gets 2 egg masses each day for a week. The next week the next aquarium in line gets 2 egg masses added daily and so on. Hence each week, 1 aquarium is set up and the oldest aquarium is ended after 7 weeks and is cleaned. This allows about 2 to 3 aquaria to have adults emerging at any one time with typically 30 to 40 males and females produced each day with an average of about 20 to 25 egg masses produced daily. Food is provided daily with food type and ration changed weekly: (1) week 1 – 10 mL/d MAFM; (2) weeks 2 and 3 – 20 mL/d MDAM; (3) weeks 4 and 5 – 40 mL/d MFFM; and (4) week 6 – no further feeding.

(5) *Collection of Adults, Mating, and Egg Collection*—Adult emergence will begin about 3 weeks after hatching at 23 °C. Once adults begin to emerge, the adults can be gently siphoned into a dry aspirator on a daily basis (for example, Fig. 4). The sex ratio of the adults should be checked to ensure that a sufficient number of males are available for mating and fertilization. One male may fertilize more than 1 female. However, a ratio of 1 male to 2 females improves fertilization. An oviposition chamber to hold emergent adults is constructed of a 5.7-L chamber with a 20-cm tall cylindrical chamber on top. The top chamber is constructed of stainless steel screen (35 mesh/2.54 cm) melted onto a plastic lid with a 17.75-cm hole. A 5-cm hole is cut into the side of the bottom chamber and a No. 11 stopper is used to close the hole. Egg masses are removed by first sliding a piece of plexiglass between the top and bottom chambers. Adult midges are then aspirated from the bottom chamber. The top chamber with plexiglass is removed from the bottom chamber and a plastic pipette is used to remove the egg masses. Egg masses are placed into a 250-mL beaker with 200 mL of test water until hatching. The top chamber is put back on top of the bottom chamber, the plexiglass is removed, and the aspirated adults are released from the aspirator into the chamber through the 5-cm hole.

(6) *Rearing Known-age Cohorts*—Egg masses are held individually in small dishes until larvae begin to emerge from the egg masses, then are transferred to clear polyethylene rectangular culture pans with a 10-cm depth and about 10 mL of <106 µm of glass beads (Sigma Aldrich, St. Louis, MO). Each pan has a 5 × 10 cm hole cut out and covered with 150 µm polypropylene mesh to allow for daily water renewal at about 7 volume additions/day using a continuous slow drip from a valve. Pans are kept at testing temperature with gentle aeration (about 2 bubble/sec). Each pan is fed 10 mL/d of *Chlorella* suspension, plus an increasing ration of Tetramin suspension beginning with 1 mL on day 4, 3 L on day 5, and 4 mL on day 6. The procedure for obtaining 3-d old larvae for life-cycle testing is described in A4.2.7.4 and A4.2.7.5.

(7) *Culture Performance Monitoring*—Feeding records and temperature recorded daily. Aeration checked daily. Records are kept on the time to first adult emergence and the success of adult emergence for each culture chamber. Reference toxicant tests are conducted periodically.

12.4.3.3 *Chironomus dilutus Culture at the U.S. Army ERDC (USACE) in Vicksburg MS:*

(1) *Culture Water*—Carbon filtered dechlorinated municipal water (Vicksburg, MS) with a hardness of about 70 mg/L as CaCO₃ and alkalinity of about 85 mg/L as CaCO₃, pH about 8.4, and with ion concentrations of about 20 mg Na/L, 2.1 mg K/L, 11 mg Ca/L, 10 mg Mg/L, 18 mg Cl/L, 8 mg SO₄/L, and <1 mg Br/L.

(2) *Foods*—Tetrafin flake fish food crushed and provided as a 400 mg/mL suspension.

(3) *Rearing Known-age Cohorts*—Cohorts for testing are reared from egg masses that are procured from a commercial supplier. Rearing of the cohort occurs in 2 phases, a hatching phase and a grow-out phase. Upon receipt from the vendor, egg masses are placed in glass 10-cm Carolina culture dishes to begin the hatching phase. Use of the small glass dish allows for monitoring of larvae development and organism transfer with a microscope for the life cycle test. Once larvae begin to emerge from the egg masses, they are transferred to rectangular polyethylene containers (45 cm × 25 cm × 15 cm) with 9 L of culture water to start the grow-out phase. A light coating of food suspension is added, which provides both food and substrate for case building; additional food is added *ad libitum* if the food supply appears insufficient. Three times a week, 50 % of the water in each grow out container is renewed. The 10-d test is started with 8-d old larvae, which routinely meet the ≤0.12 mg AFDW standard. For chronic life cycle testing with up to 3-d old larvae, larvae are maintained in the culture bowl used for hatching; after larvae begin to emerge from the egg masses, a light coating of food is added to the culture bowl.

12.4.3.4 *Chironomus dilutus Culture at Smithers Viscient, LLC, Wareham MA:*

(1) *Culture Water*—Well water with a hardness of 60 to 80 mg/L as CaCO₃ and alkalinity of about 20 mg/L as CaCO₃, pH 7.0, and conductivity of 400 to 600 µS/cm.

(2) *Foods*—Midge cultures are fed suspensions of flaked fish food (Sera Vipran Staple) prepared in well water at varying concentrations, depending on life-stage and holding conditions.

(3) *Substrate*—Primary cultures are maintained with a substrate of blended cellulose filter paper. Larvae reared to the 3rd instar for 10-d testing are maintained with a substrate of fine silica sand and finely ground peat moss. Larvae reared to the 1st instar for life-cycle testing are maintained with a substrate of green algae, *Ankistrodesmus falcatus*.

(4) *Primary Culture*—The primary culture is maintained in several 57-L aquaria. Cellulose filter paper is blended with well water and placed in each aquaria at a depth of 1 to 2 cm. The overall water depth in each aquaria is maintained at 15 cm, and each aquaria is covered with fine mesh to retain emerged adults. Aquaria are maintained under flow through conditions (about 5 volume additions/d) along with constant aeration from an air stone. Each aquarium is initially

from about 8 egg masses. Additional larvae may be added a week later to provide supplemental emergence during the peak reproductive period. Each aquaria of the primary culture is provided a suspension of flaked fish food (100 mg/mL) daily. The volume provided is between 5 and 10 mL, depending on the size and age of the larval population.

(5) *Collection of Adults, Mating, and Egg Collection*—Emerged adults are kept within the primary aquaria from which they were reared and allowed to swarm. Egg masses produced are collected from the aquaria on a daily basis. In instances when the swarm is lacking density and egg mass production is poor, several males and females (10 to 15 each) may be removed and placed in a large flask (for example, 2 L) to facilitate mating. The flask contains a small amount of water (1 cm depth) for egg masses to be deposited, and is covered with a fine mesh net to retain the adults. Adults are supplemented or replaced daily as needed until a desired number of egg masses have been collected.

(6) *Rearing Known-age Cohorts*—Egg masses are collected and placed in 250 mL crystalizing dishes with well water and are observed daily until hatch. To provide 1st instar larvae (1 to 4 days old) for life-cycle testing, hatched larvae of 2 egg masses are transferred to a shallow flat-bottomed glass 1-L bowl containing 0.8 L of well water. *Ankistrodesmus falcatus*, a unicellular green algae, is added as a substrate to facilitate tube building. An aliquot of 10 mg of flaked fish food delivered as a suspension is provided daily. To provide 3rd instar larvae for 10-d testing, hatched larvae of 2 egg masses are transferred to an 11-L plastic tub containing about 8 L of well water and a thin (1 to 2 mm) layer of fine silica sand mixed with finely ground peat moss. *Ankistrodesmus falcatus* is typically added on the first day to facilitate tube building. Larvae are fed 40 mg of flaked fish food delivered as a suspension daily. This ration is typically doubled following 3 days of rearing to provide third instar larvae within the expected period of time (9 to 10 days). All known-age rearing is performed under static conditions with continuous aeration.

(7) *Culture Performance Monitoring*—Feeding records and temperature are recorded daily. Dissolved oxygen and pH are measured weekly in the primary culture, and every other day during the rearing of known-age larvae for testing. Primary cultures are observed for time to first emergence. Reference toxicant tests are conducted periodically with potassium chloride.

12.4.3.5 *Chironomus dilutus* Culture at the EAG Laboratories-Easton, MD:

(1) *Culture Water*—Water used for holding the test organisms is obtained from a well located on the EAG Laboratories site in Easton, Maryland. The water is sand filtered, treated to remove any particulate matter over 0.45 µm particle size and UV sterilized before use. The water typically has hardness of about 140 mg/L as CaCO₃ and alkalinity of 180 mg/L as CaCO₃. Specific Conductance of about 365 µS/cm and total organic carbon of less than 2 mg/L. These measurements are taken weekly to monitor the consistency of the well water. The water is characterized as moderately hard.

(2) *Foods*—10 g/L *Chlorella*, invertebrate slurry (Ingredients: Otohime A, Tetramin flake, alfalfa leaf powder, vitamin mix, Algamac Protein Plus, dried spirulina powder (organic), *Chlorella*, kelp powder (organic), natural astaxanthin) 15 g/L Tetramin® flake food suspension in reverse osmosis water.

(3) *Substrate*—Culture of 1 to 2 cm of clean silica sand.

(4) *Adult Culture*—*Chironomus dilutus* cultures are maintained in 20-L glass aquaria containing about 8 L of well water under flow-through conditions and 1 to 2 cm of clean silica sand as a substrate. The flows to each aquarium are checked at least weekly and are adjusted as needed. Two egg masses are initially placed either directly on top of the sand or into a 600-mL glass beaker (containing water only) within the aquaria. When using the glass beaker, a half egg mass is placed into each beaker until the egg mass hatches. The newly hatched larvae are then gently rinsed into the tank following the hatch. A new tank with a new egg mass will be started within a few days of each other in order to stagger the emergence of the adult midges and to have a continuous supply of males and females to be able to produce sufficient egg masses required for testing. Emerged midges are collected daily during the emergence period into a plastic cylindrical tube with rubber stoppers on either end using a vacuum pump and released into breeding cages (covered aquaria containing about 4-L of UV sterilized well water). Egg masses are collected daily and are either held until the hatched organisms are appropriately aged for testing or used to continue the culture population.

(5) *Typical Feeding in Culture*—No food is provided to egg masses. Week 1 (after hatch) 1.5 mL of 10 g/L *Chlorella*/beaker/d or 5 mL of invertebrate slurry. Week 2 to 3: 10 mL of invertebrate slurry/aquaria. Week 4 to 5: 20 mL of invertebrate slurry/aquaria.

(6) *Obtaining and Rearing Known-age Cohorts*—Appropriate number of egg masses are collected from the culture. Half of an egg mass or 1 egg mass is placed into 600-mL glass beaker containing about 500 mL of UV sterilized well water. The beakers are placed into 40-L aquaria containing water equipped with aquarium heaters to maintain required temperature. Egg masses are observed daily, and date of the hatch is recorded. At this time, any unhatched egg masses are removed and discarded. When the organisms are 2 to 3 d old, they are used to start a test or gently rinsed into a 40-L aquaria containing 16 to 20 L of UV sterilized well water. The larvae are reared until reaching an appropriate age as needed to be used in a test or are used to continue the culture population.

(7) *Typical Feeding During Holding for Tests:*

(a) *Rearing to 10 days*—Day 0 to 3—1.25 mL of 10 g/L *Chlorella*/beaker/d. Day 3 to 7: 2.5 mL of invertebrate slurry/tank/d until Day 7 (after organisms are transferred into culture aquaria with sand substrate). Day 8 to 9—5 mL slurry/tank/d.

(b) *Rearing to 4 days*—For 4-d rearing, organisms are in 600 mL beakers for the entire duration of holding. Feeding regime is—Day 0 to 4: 1.25 mL of 10 g/L *Chlorella*/beaker/d; Day 3 to 4: 0.25 mL of 15 g/L Tetramin® flake food suspension.

12.5 *Methods for Preparing Foods Used in Testing and Culture:*



Documents

Document Sharing Hub

12.5.1 Preparation of Flake Fish Food Suspensions—In the previous version of this standard, flake fish food suspension was the recommended diet for testing with midges ([Annex A2](#) and [Annex A4](#)), prepared by high energy blending of flake fish food with water, resulting in a very fine suspension. Subsequent studies at U.S. EPA-Duluth has indicated that this preparation is unnecessarily fine, and is a less efficient means to deliver food to sediment test organisms. This version of the standard recommends a suspension of fine flakes, which is suspended only by swirling crushed flakes in water. This fine flake suspension has also been added as part of the feeding in tests with amphipods ([Annex A1](#) and [Annex A3](#)). However, the blended flake food suspension is still used for the first few days of the life-cycle test with midges ([Annex A4](#)). Procedures for preparing both types of flake food suspension are provided here.

12.5.1.1 Preparation of Fine Flake Suspension—The flake food suspension is prepared from dry fish food flakes (Tetramin® Tropical Flakes or equivalent; Tetra Holding (USA), Inc., Blacksburg, VA). Dry flakes are crushed through a U.S. standard (#50) sieve (300-µm opening), then apportioned by weight and added to a measured amount of deionized water. Feeding rates used in the amphipod and midge test methods are specified by weight; the concentration of fine flakes in the suspension can be adjusted so that the proper ration is delivered in a convenient volume (for example, 1 mL) by multiplying the target daily ration by the volume of food prepared (in mL). For example, 100 mL of a food solution delivering 0.25 mg/mL can be prepared by adding $0.25 \text{ mg/mL} \times 100 \text{ mL} = 25 \text{ mg}$ flakes to 100 mL of deionized water. Other concentrations can be made to allow feeding of the desired ration in a desired volume. The flake food solution is gently agitated to wet and distribute the flakes (not blended) and stored under refrigeration; fresh solution should be prepared every 3 days (that is, fresh solution is used the day of preparation and for 2 additional days).

12.5.1.2 Preparation of Blended Suspension—While most uses of a flake fish food suspension in this standard are for the fine flake suspension described in [12.5.1.1](#), a blended suspension is used for the first 3 days of the midge life cycle test ([Annex A4](#)). The blended suspension is prepared from the same dry fish food flakes (Tetramin® Tropical Flakes or equivalent; Tetra Holding (USA), Inc., Blacksburg, VA). Dry flakes are combined with de-ionized water and blended at high speed for a minute to produce a finely divided suspension. Four grams of flakes per L of water provides a solution that will deliver 6 mg of dry food in 1.5 mL of solution; other proportions and delivery volumes may be used to provide the desired ration. Store the blended solution under refrigeration, and mix thoroughly before use. Store no longer than 2 weeks.

12.5.2 Preparation of Diatom Suspension—In the current version of this standard, a suspension of diatoms is added as component of the diet for testing with amphipods (1 of 2 alternatives; [Annex A1](#) and [Annex A3](#)). While initial studies evaluating a diatom diet for testing with *H. azteca* used cultured diatoms (*Mayamaea* sp. and *Nitzschia* sp.; Soucek et al. 2106), commercially available diatom solutions may be a more convenient source for many laboratories. Preparation of a

diatom solution from a commercially available solution is described in this subsection; other sources of diatom solution may also be successful, though their adequacy for use in testing should be verified, and the preparation procedures may require modification from those listed here.

12.5.2.1 Diatom Source—The diatom, *Conticribra weissflogii* (WoRMS (2015) ([169](#)); formerly known as *Thalassiosira weissflogii*); [12.5.2.1](#), is a unicellular microalgae found in marine, brackish, and freshwater environments. A stabilized concentrate sold as *T. weissflogii* is available commercially (for example, Reed Mariculture, Inc, Campbell, CA Instant Algae® TW1200). The shelf life of the concentrate is 12 to 14 weeks under refrigeration.

12.5.2.2 Preparation of Diatom Food Solution—The *C. weissflogii* concentrate is in a saline solution and requires washing to reduce the salt content; this can be accomplished by adding fresh water to aliquots of concentrate in a 4:1 ratio, mixing, then centrifuging to re-concentrate the diatoms. This process is repeated, then conductivity is used to verify that ion concentrations have been reduced. The specifics of this process can be adjusted to the equipment available in a particular laboratory; the following steps provide 1 example using 100-mL centrifuge tubes.

(1) Combine about 20 mL of diatom concentrate and about 80 mL of clean control or culture water in each of 8 100-mL centrifuge tubes and mix well. Adding water using a squirt bottle may help suspend the diatoms which may stick to the side of the tube.

(2) After mixing, centrifuge the tubes for about 15 min at about $1700 \times g$ (other centrifugation durations and intensities may be suitable also). Decant the supernatant and the add water to the same total volume used initially, mix well, and repeat the centrifugation process.

(3) After the second centrifugation, decant the supernatant and transfer the concentrated diatoms to a single container and add water to bring the diatom solution to about three times the original volume of diatoms that were rinsed and mix (in this example $20 \text{ mL concentration} \times 8 \text{ tubes} \times 3 = 480 \text{ mL}$).

(4) Check the conductivity of the final suspension; it is typically about $700 \mu\text{S/cm}$; if it is substantially higher than that, another centrifugation cycle can be used to reduce conductivity further.

(5) Store the washed concentrate under refrigeration; the shelf life after washing is about 2 weeks.

12.5.2.3 Adjusting Solids Content and Calculating Ration Volumes—The solids content of the washed diatom solution is adjusted to a consistent value to allow delivery of known rations. The selection of solids content and delivery volumes is arbitrary, though the example described results in delivery volumes that are easily delivered with most laboratory pipettes; other concentrations and volumes may be used as long as they deliver the appropriate ration. First, the solids concentration of the washed diatoms solution is determined by placing measured volumes (for example, 3 mL) into triplicate tared weigh pans, drying for 24 h at 70°C , cooling the pans, then determining the mass of dry solids. Based on this value, add water is as needed to the washed diatom solution to bring the

5 mg solids/mL. This solution can then be used to deliver 0.5 mg in 100 μ L or 0.75 mg in 150 μ L.

12.5.3 Preparation of YCT—YCT is a blend of several ingredients; the name derives from its original components, yeast, Cerophyl® (a commercial preparation of cereal grass leaves), and trout chow. As described in this subsection, it is prepared by combining equal proportions of solutions prepared with each of the three ingredients.

12.5.3.1 Preparation of Incubated Trout Chow—The trout chow solution is incubated at room temperature for a week, so it must be prepared ahead of the other 2 solutions. Trout chow (for example, $\frac{1}{8}$ in. pellets) should meet current U.S. Fish and Wildlife Service specifications, and is available from many suppliers including Ziegler Bros., Inc., (400 Gardners Station Road, Gardners, PA 17324 USA (717) 677-6181) and Skretting USA (712 East 2400 North, Tooele, UT 84074, (800) 521-9092). Add 5.0 g of trout chow pellets to 1 L of deionized water. Mix well in a blender and pour into a 2-L separatory funnel or similar container. Aerate the solution and allow to incubate at room temperature for a week. Because of the unpleasant odor produced, incubating in a hood or other ventilated space is recommended. After a week of incubation, allow the solution to settle for 1 h, then filter the supernatant through a fine mesh screen (for example, 110 μ m mesh).

12.5.3.2 Preparation of Yeast Solution—Add 5.0 g of dry baker's yeast to 1 L of deionized water and mix (by hand shaking, magnetic stirrer, or low speed blending) until the yeast granules are completely dispersed.

12.5.3.3 Preparation of Cereal or Alfalfa Leaf Solution—Cerophyl® was a preparation of wheat grass sold as a nutritional supplement that is no longer available. In its place, dried, powdered wheat grass, "cereal leaves," or alfalfa leaves are used. These are widely available in health food stores and other sources; certified organic sources may be less likely to contain undesirable contaminants. To prepare the cereal/alfalfa leaf solution, combine 5.0 g of powdered dried cereal/alfalfa leaves with 1 L of de-ionized water and either blend on high speed for 5 minutes, or stir overnight using a magnetic stirrer. Allow the solution to settle for at least 1 h, then decant the overlying solution leaving behind any settled solids.

12.5.3.4 Combine Solutions and Adjust Solids Content—YCT is prepared by combining equal volumes of the three solutions and mixing well. The combined solution should contain 1.7 to 1.9 g solids per L, which can be verified by drying a small aliquot and weighing the residual. Water can be added as needed to reach the desired solids content. YCT should be kept refrigerated, and used for up to 2 weeks. A large batch can be portioned into small plastic bottles (for example, 50 to 100 ml) and frozen for up to three months before use.

13. Interpretation of Results and Reporting

13.1 Interpretation of Results:

13.1.1 Quality Assurance Project Plan—While data analysis is conducted following completion of a test, the methods by which data will be analyzed should be considered before conduct of the study, as the means by which data will be analyzed will affect study design. These considerations are part of project planning and should be included in the quality assurance project plan (11.1.1).

13.1.2 Scope—Selecting appropriate statistical analysis procedures can be a complex issue, and it is dependent on a large number of factors, many of which may be study-specific. Because the test methods described in this standard may be applied in studies with different goals, it is not possible to specify statistical procedures that will be universally appropriate. This section provides some suggested analyses that are likely to apply in many cases, together with additional discussion that may be helpful in determining when alternatives should be considered. However, nothing in this section should be taken as direction to apply particular statistical procedures if the characteristics of the data set or the goals of the study suggest that they may be inappropriate.

13.1.3 Sources for Additional Information—There are many sources of more detailed guidance on the application of statistical tests to ecotoxicity data, including discussion of when particular procedures may be more or less applicable, and alternatives that can be used to evaluate particular hypotheses or for data with different characteristics. Both OECD (2006) (16) and Environment Canada (2007) (11) have published such guidance, in addition to more generalized information found in sources such as Ellersieck and LaPoint (1995) (170), Newman (1995) (171), Bretz et al. (2011) (172), Practice E1847, and primary sources cited therein.

13.1.4 Software for Statistical Analysis—Previous versions of this standard have provided detailed instructions for de novo calculation of various parameters and test statistics. Computer software capable of conducting these analyses is now widely available, so these details have been omitted from this version of the standard. Software packages such as R, Systat, SAS, Jump, Matlab, Statistica, SPSS, and Minitab can perform analyses (or close analogs) discussed in this section.

13.2 Sediment Testing Scenarios—Sediment tests are conducted to determine whether contaminants in sediment are harmful to benthic organisms. Sediment tests are commonly used in studies designed to: (1) evaluate hazards of dredged material, (2) assess site contamination in the environment (for example, to rank areas for cleanup), and (3) determine effects of specific contaminants, or combinations of contaminants, through the use of sediment spiking techniques. Each of these broad study designs has specific statistical design and analytical considerations, which are described as follows.

13.2.1 Dredged Material Hazard Evaluation—In these studies, *n* (number) sites are compared individually to a reference sediment. The statistical procedures appropriate for these studies are generally pairwise comparisons. Additional information on toxicity testing of dredged material and analysis of data from dredged material hazard evaluations is available in USEPA-USACE (1998) (91).

13.2.2 Site Assessment of Field Contamination—Surveys of sediment toxicity are often included in more comprehensive analyses of biological, chemical, geological, and hydrographic data. Statistical correlation can be improved and costs may be reduced if subsamples are taken simultaneously for sediment tests, chemical analyses, and benthic community structure determinations. There are several statistical approaches to field assessments, each with a specific purpose. If the objective is to compare the response or residue level at the test site to that of

a control sediment, then the pairwise comparison approach described as follows is appropriate. If the objective is to compare among all sites in the study area, then a multiple comparison procedure that employs an experiment-wise error rate is appropriate. If the objective is to compare among groups of sites, then orthogonal contrasts are a useful data analysis technique.

13.2.3 Sediment Spiking Experiments—Sediments spiked with known concentrations of chemicals can be used to establish cause and effect relationships between chemicals and biological responses. Results of toxicity tests with test materials spiked into sediments at different concentrations may be reported in terms of as lethal percent concentrations (for example, LC10, LC20, LC50), effect percent concentrations (for example, EC10, E20, EC50), inhibition percent concentrations (for example, IC10, IC20, IC50), or no-observed-effect concentrations or lowest-observed effect concentrations (NOEC, LOEC). The statistical approach for spiked sediment toxicity tests also applies to the analysis of data from water-only reference-toxicity tests.

13.3 Experimental Design—The objectives of a toxicity test are to quantify contaminant effects on test organisms exposed to water-borne contaminants, natural or spiked sediments or dredged materials and to determine whether these effects are statistically different from those occurring in a control or reference sediment or a control water treatment. Each experiment consists of at least 2 treatments: the control and one or more test treatment(s). The test treatment(s) in sediment exposures consist(s) of the contaminated or potentially contaminated sediment(s).

13.3.1 Control Sediment—Sediment toxicity tests include a control sediment (sometimes called a negative control). A control sediment is a sediment that is essentially free of contaminants, is used routinely to assess the acceptability of a test, and is not necessarily collected near the site of concern. Any contaminants in control sediment are thought to originate from the global spread of pollutants and do not reflect any substantial input from local or nonpoint sources. A control sediment provides a measure of test acceptability, evidence of test organism health, and a basis for interpreting data obtained from the test sediments (3.3.5). A reference sediment is typically collected near an area of concern (for example, a disposal site) and is used to assess sediment conditions exclusive of material(s) of interest (3.3.15). Testing a reference sediment provides a site-specific basis for evaluating toxicity. Comparisons of test sediments to multiple control or reference sediments representative of the physical characteristics of the test sediment (that is, grain size, organic carbon) may be useful in these evaluations). Section 13.5 provides guidance on using hypothesis testing to statistically determine significant differences between test sediments and reference sediments or control sediments (for example, by analysis of variance and mean separation).

13.3.1.1 Test Acceptability—In general, the performance of test organisms in the negative control is used to judge the acceptability of a test (Table A1.1, Table A2.1, Table A3.1, Table A4.1, Table A5.1, Table A6.1, Table A7.1, Table A8.1), and either the negative control or reference sediment may be

used to evaluate performance in the experimental treatments, depending on the purpose of the study. Any study in which organisms in the negative control do not meet test acceptability criteria should be considered questionable because it suggests that adverse factors affected the test organisms. Key to avoiding this situation is using only control sediments that have a demonstrated record of performance using the same test procedure. This includes testing of new collections from sediment sources that have previously provided suitable control sediment. Because of the uncertainties introduced by poor performance in the negative control, such studies should be repeated to ensure accurate results. However, the scope or sampling associated with some studies may make it difficult or impossible to repeat a study. Some researchers have reported cases where performance in the negative control is poor, but test acceptability criteria are met in a reference sediment included in the study design. In these cases, it might be reasonable to infer that other samples that show good performance are probably not toxic; however, any samples showing poor performance should not be judged to have shown toxicity, since it is unknown whether the adverse factors that caused poor control performance might have also caused poor performance in the test treatments.

13.3.1.2 Use of Sand Control to Evaluate the Quality of the Water or Diet Used in Toxicity Testing—The water and diet used to conduct a water or sediment toxicity test should be sufficient to support test organisms independently from the quality of the substrate tested (for example, test acceptability criteria controls in individual tests outlined in Table A1.1, Table A2.1, Table A3.1, Table A4.1, Table A5.1, Table A6.1, Table A7.1, and Table A8.1. If test organisms grow substantially better in field sediments than in a substrate like sand, then it suggests that the test organisms are likely relying on sediments to provide for a meaningful portion of their needs. This in turn means that differences in performance of a test organism among sediments could be confounded by the “nutritional quality” of the sediment, separate from the effects of sediment contaminants. Control sediment(s) should be sediments that readily support survival, growth and reproduction of test organisms, but are not so rich that growth or reproduction of test organisms in a control sediment is consistently better than reference sediments.

13.3.1.3 Periodic Testing of Controls—Periodic testing of both a regular control sediment and a sand control is recommended (perhaps testing both at least once/year). If survival, weight, biomass, or reproduction of test organisms in the regular control sediment exceeds that in the sand control by more than about 20 to 30 %, the testing laboratory should consider an alternate control sediment. This is to avoid an unrealistic expectation of performance that might, among other things, cause uncontaminated, but less “nutritious” sediments to appear toxic. One source of the sand is GRANUSIL® mineral filler, such as GRANUSIL® #5010, Unimin Corporation, New Canaan, CT (GRANUSIL #5010 is a mixture of various size silica sand, such that the percent retained on mesh according to Specification E1706-20 is 100 %).

sieve; 22.5 % on No. 40 sieve; 54.7 % on No. 50 sieve; 17.4 % on No. 70 sieve; 4.2 % on No. 100 sieve, and 0.9 % on No. 140 sieve).

13.3.1.4 Physico-chemical Characteristics of Sediment—Natural physico-chemical characteristics of sediment such as sediment texture may influence the response of test organisms (DeWitt et al. 1988) (173). The physico-chemical characteristics of test sediment should be within the tolerance limits of the test organism. Ideally, the limits of a test organism should be determined in advance; however, controls for factors such as grain size and organic carbon can be evaluated if the recommended limits are approached or exceeded in a test sediment. See [Annex A1](#) to [Annex A8](#) for information on physicochemical requirements of test organisms. If the physicochemical characteristics of a test sediment exceed the tolerance limits of the test organism, it may be desirable to include a control sediment that encompasses those characteristics. The effects of some sediment characteristics (for example, grain size or total organic carbon) on sediment test results may be addressed with regression equations (DeWitt et al. 1988 (173); Ankley et al. 1994b (45)). The use of formulated sediment can also be used to evaluate physico-chemical characteristics of sediment on test organisms (for example, Section 7.2 in USEPA 2000 (8)).

13.3.2 Establishing Toxicity of Test Sediments Relative to Reference Sediment(s)—Reference sediments are defined as a whole sediment near an area of concern used to assess sediment conditions exclusive of material(s) of interest (3.3.15). Reference sediments may be used as an indicator of localized sediment conditions exclusive of the specific pollutant input of concern. Such sediments would be collected near the site of concern and would represent the background conditions resulting from any localized pollutant inputs as well as global pollutant input. This is the manner in which reference sediments are used in dredged material evaluations. Toxicity of test sediments can be established based on comparisons to a single reference sediment or to multiple reference sediments. Ideally, the reference sediment(s) would be tested in the same batch of sediments with the test sediments and with any control sediments(s). An advantage of establishing toxicity of test sediments (site sediments) relative to reference sediments is that the reference sediments may better represent the range of physical characteristics of the test sediments at a site. However, it may be difficult to identify reference sediments near a site of concern that are representative of site sediments (for example, purported reference sediments may be contaminated with known or unknown materials what influence the response of test organisms).

13.3.2.1 Reference Envelope—A reference-envelope approach has been used to classify site sediments as toxic or not toxic based on a statistical representations of the distributions of toxicity data and chemistry data in site sediments relative to reference sediments (for example, Hunt et al. 2001 (174); Ingersoll et al. 2009 (175), 2014 (100); Kemble et al. 2013 (176); Wang et al. 2013 (87); Besser et al. 2015 (84)). The reference-envelope approach is a procedure for assessing sediment toxicity that was developed to overcome the limitations associated with the use of single control sediments or multiple control sediments for this purpose. Reference sedi-

ment can be used to account for differences in the non-contaminant characteristics of test sediments (for example, grain size, total organic carbon) and for overcoming the low statistical power associated with comparing multiple test sediments to a single control sediment or to a single reference sediment with physical characteristic that may differ from site sediments. Application of the approach necessitates identification of reference sediment samples for each toxicity test endpoint that is evaluated.

13.3.2.2 Example Application of Reference Envelope Approach—An approach described by Ingersoll et al. (2014)(100) is provided below as an example of how site sediments can be identified as reference sediments using chemical selection criteria and biological selection criteria and how toxicity can be evaluated in the site sediments relative the distribution of chemistry or biological endpoint determined for the reference sediments.

(1) As a first step, site samples with chemical characteristics representative of reference conditions are identified (for example, samples substantially free of contamination). For example, Ingersoll et al. (2014) (100) used the following chemical selection criteria to select reference sediments based on site sediments with: (1) Mean probable effect concentration quotients (PECQ; MacDonald et al. 2000 (26), Ingersoll et al. 2001 (27)) less than 0.1 based on measured concentrations of organic contaminants in whole sediment (including polycyclic aromatic hydrocarbons, polychlorinated biphenyls, and pesticides) and based on measured concentrations of metals in whole sediment and with (2) Chronic sum equilibrium-partitioning sediment benchmark toxic units (Σ ESBTU) less than 0.1 (USEPA 2003) (23).

(2) As a second step, site samples that meet test acceptability criteria for control survival, weight, or reproduction are identified. For example, Ingersoll et al. (2014) (100) used the biological selection criteria to select reference sediments based on site samples that met test acceptability criteria for control sediment (for example, test acceptability criteria for control sediment summarized in [Table A2.1](#) and [Table A4.1](#)).

(3) As a third step, if an analysis of variance for a particular toxicity endpoint (for example, 28-d weight of *H. azteca*), within a batch of samples was not significant ($p > 0.05$), data for that endpoint might not be further evaluated (for example, Besser et al. 2015 (84)).

(4) Site samples that met both the chemical selection criteria and biological selection criteria were included in the pool of reference sediments (Ingersoll et al. 2014) (100). Following the identification of reference sediment samples, the range of the biological responses in the test sediment relative to the reference sediment can be determined for each toxicity test and for each toxicity endpoint measured.

(5) For example, Ingersoll et al. (2014) (100) defined a reference envelope as the range of biological responses that encompassed 100 % of the response data for the reference sediment samples. Accordingly, the lower limit of the reference envelope was calculated as the minimum control-adjusted response value for each toxicity test and endpoint, using the data for the reference sediment samples that were selected for each toxicity test. The reference envelope was then used to

define the normal range of responses associated with exposure of toxicity test organisms to relatively uncontaminated sediment samples. Sediment samples with effect values lower than the lower limit of the normal range of control-adjusted responses for the reference sediments were designated as toxic for the endpoint under consideration. The sediment samples also were designated as toxic or not toxic based on the results of multiple endpoints for each toxicity test (for example, survival, weight, biomass, or reproduction). Finally, sediment samples were designated as toxic or not toxic based on the results obtained from any of the toxicity test endpoints.

(6) Whereas, classification of sediment samples as toxic or not toxic provides important information for assessing site-specific sediment quality conditions, additional information on the magnitude of toxicity can contribute to such evaluations. For this reason, toxic sediment samples were further classified to identify moderately toxic and highly toxic sediment samples (Ingersoll et al. 2014) (100). Highly toxic (HT) sediment samples were identified based on a greater than 10 % reduction in survival, weight, biomass, or reproduction relative to the lower limit of the reference envelope. Moderately toxic (MT) sediment samples were identified based on survival, weight, biomass, emergence, or reproduction that fell less than 10 % below the lower limit of the reference envelope.

13.4 *Evaluating Relationships Between Sediment Toxicity and Sediment Chemistry:*

13.4.1 Sediments spiked with known concentrations of chemicals can be used to establish cause and effect relationships between chemicals and biological responses. Results of toxicity tests with test materials spiked into sediments at different concentrations may be reported in terms of lethal concentrations (for example, LC50), effect concentrations (for example, EC50, inhibition concentrations (for example, IC50), or as an NOEC (no observed effect concentration) or LOEC (lowest observed effect concentration). Most studies with spiked sediment are often started only a few days after the chemical has been added to the sediment. Consistent spiking procedures should be followed in order to make interlaboratory comparisons (10.6).

13.4.2 Evaluating effect concentrations for chemicals in sediment requires knowledge of factors controlling the bioavailability. Similar concentrations of a chemical in units of mass of chemical per mass of sediment dry weight often exhibit a range in toxicity in different sediments (Di Toro et al. 1990 (4), 1991 (2)). Effect concentrations of chemicals in sediment have been correlated to interstitial water concentrations, and effect concentrations in interstitial water are often similar to effect concentrations in water-only exposures. The bioavailability of nonionic organic compounds are often inversely correlated with the organic carbon concentration of the sediment. Whatever the route of exposure, the correlations of effect concentrations to interstitial water concentrations indicate predicted or measured concentrations in interstitial water can be useful for quantifying the exposure concentration to an organism. Therefore, information on partitioning of chemicals between solid and liquid phases of sediment may be useful for establishing effect concentrations.

13.4.3 Toxic units can be used to help interpret the response of organisms to multiple chemicals in sediment. A toxic unit is the concentration of a chemical divided by an effect concentration. For example, a toxic unit of exposure can be calculated by dividing the measured concentration of a chemical in pore water by the water-only LC50 for the same chemical (Ankely et al. 1991) (177). Toxic units could also be calculated by dividing the concentration in a whole sediment sample by a threshold concentration in whole sediment (Long and Morgan 1991 (178), MacDonald et al. 2000 (26)). Toxicity expressed as toxic units may be summed and this may provide information on the toxicity of chemical mixtures (Ingersoll et al. 2005b) (179).

13.4.4 Field surveys can be designed to provide either a qualitative reconnaissance of the distribution of sediment contamination or a quantitative statistical comparison of contamination among sites. Surveys of sediment toxicity are usually part of more comprehensive analyses of biological, chemical, geological, and hydrographic data (USEPA 2002a, b, and c (20, 21, 22), Table 1). Statistical correlation can be improved and costs reduced if subsamples are taken simultaneously for sediment tests, chemical analyses, and benthic community structure.

13.4.5 Descriptive methods such as toxicity tests with field-collected sediment should not be used alone to evaluate sediment contamination. An integration of several methods using the weight of evidence is needed to assess the effects of contaminants associated with sediment (Long et al. 1990 (180), Ingersoll et al. 1997 (181), MacDonald et al. 2000 (26), USEPA 2002a, b, and c (20, 21, 22), Wenning et al. 2005 (19)). Hazard evaluations integrating data from laboratory exposures, chemical analyses, and benthic community assessments (the Sediment Quality Triad) provide strong complementary evidence of the degree of pollution-induced degradation in aquatic communities (Canfield et al. 1994 (182), 1996 (183), 1998 (184); Chapman et al. 1997 (29); USEPA 2002abc (20, 21, 22)).

13.4.6 Toxicity Identification Evaluation (TIE) procedures can be used to help provide insights as to specific contaminants responsible for toxicity in sediment (USEPA 2007) (6). For example, the toxicity of contaminants such as metals, ammonia, hydrogen sulfide, and nonionic organic compounds can be identified using TIE procedures.

13.5 *Approaches for Statistical Analysis of Data:*

13.5.1 *General Analysis Approaches for Sediment Toxicity Tests*—In most instances, results from sediment toxicity tests are analyzed by one of two approaches, either: (1) hypothesis-based or (2) regression. Hypothesis-based analysis compares the performance of organisms among two or more treatments to determine whether statistically significant differences in the level of performance exist. In its simplest form, this could be determining whether performance in a potentially contaminated sediment is different (for example, lower) than in a control sediment. In many cases, several sediments are compared to one or more control or reference treatments. The output of this analysis is simply which sediments are statistically separable from others. Hypothesis-based analysis is often used in evaluation of dredged material. Documents

1998 (91)) or in surveying a potentially contaminated site for sediment toxicity. It can also be used to determine whether performance is significantly different from a specified level of performance (for example, is the observed survival significantly lower than 90 %). Hypothesis-based analysis carries no inference as to why performance differs, only that the outcomes are statistically separable. The goal of regression analysis is to evaluate the overall relationship between the concentration of contaminant in sediment and the level of organism performance, often referred to as an exposure-response relationship. Regression approaches are typically used to establish effect concentrations, such as the median lethal concentration (LC50), or exposures associated with specific levels of effect, such as a 20 % reduction in growth or reproduction (often referred to as the EC20 or 20 % effect concentrations). Effect concentrations may be established for any endpoint or any level of effect. In sediment toxicity testing, regression analysis is most commonly used in spiked sediment tests where the test design involves a structured series of exposure concentrations, though it is also sometimes used to evaluate exposure-response relationships in field sediments where gradients of chemicals or characteristics causing toxicity are known or suspected (for example, Ingersoll et al. 2014 (100)).

13.5.2 Experimental Unit—For sediment toxicity testing, each test chamber to which a single application of treatment is applied is an experimental unit. The important concept is that the treatment (sediment) is applied to each experimental unit as a discrete unit. Experimental units should be independent and should not differ systematically. Although some measurements (for example, survival, reproduction in the midlife-cycle test) are made on individual test organisms, for statistical comparison, measurements made on individuals are still aggregated at the level of the test chamber.

13.5.3 Randomization—Randomization is the unbiased assignment of treatments within a test system, and to the exposure chambers, to ensure that no treatment is favored and that observations are independent. Locations of exposure chambers should be randomized. Impartial assignment of organisms to exposure chambers is discussed in the individual test methods described in the Annexes. It is generally impractical to apply fully randomized assignment of individual

organisms to test chambers; instead, stocking organisms into test chambers using stratified random assignment is more common.

13.5.4 Replication—Replication is the assignment of a treatment to more than one experimental unit. The variation among replicates is a measure that provides an estimate of within-treatment error for assessing the significance of observed differences between treatments. General guidance on replication is given within the individual test methods described in the Annexes, but the number of replicates within a given test should be determined on the basis of the needs of the test. These needs may differ based on, among other things, the data analysis approach to be used (that is, hypothesis-based vs. regression approaches).

13.5.5 Considerations for Replication Using Hypothesis-based Analysis—For hypothesis-based analysis, the number of replicates is directly related to the ability to detect differences of a given magnitude (the minimum detectable difference or MDD). These can be understood in the context of power analysis, as described in 13.5.5.1 to 13.5.5.7).

13.5.5.1 Minimum Detectable Difference (MDD)—As the minimum difference between treatments which the test is required or designed to detect decreases, the number of replicates required to meet a given significance level and power increases. No consensus currently exists on what constitutes a biologically meaningful change in organism performance. Therefore, the appropriate statistical minimum significant difference should be a data quality objective (DQO) established as part of study planning, with consideration of the study goals and associated data requirements, the logistics and economics of test design, and the ultimate use of the sediment toxicity test results.

13.5.5.2 Power Analysis for Hypothesis Testing—The prototypical hypothesis evaluated using a sediment toxicity test is whether the biological response to a treatment sample differs from the response to the control or reference sample(s). Fig. 5 presents the possible outcomes and decisions that can be reached in a statistical test of whether performance in a treatment is lower than in the control or reference sample(s). The null hypothesis is that no difference exists between the mean control samples(s) or the mean reference sample(s) and the mean treatment responses. The alternative hypothesis is

		True Condition	
		Treatment = Control/Reference	Treatment < Control/Reference
Hypothesis Test Outcome	Null Hypothesis Treatment = Control/Reference	Correct $1 - \alpha$	Type II Error β
	Alternative Hypothesis Treatment < Control/Reference	Type I Error α	Correct $1 - \beta$ (power)

NOTE 1—Alpha (α) represents the probability of making a Type I error (false positive); beta (β) represents the probability of making a Type II error (false negative). Power is $1 - \beta$, the chance of correctly determining that a difference exists.

FIG. 5 Contingency Table Showing Relationship Among Error Types, Significance Level, and Power

that the treatment has lower performance to the control sample(s) or reference sample(s) (indicating toxicity).

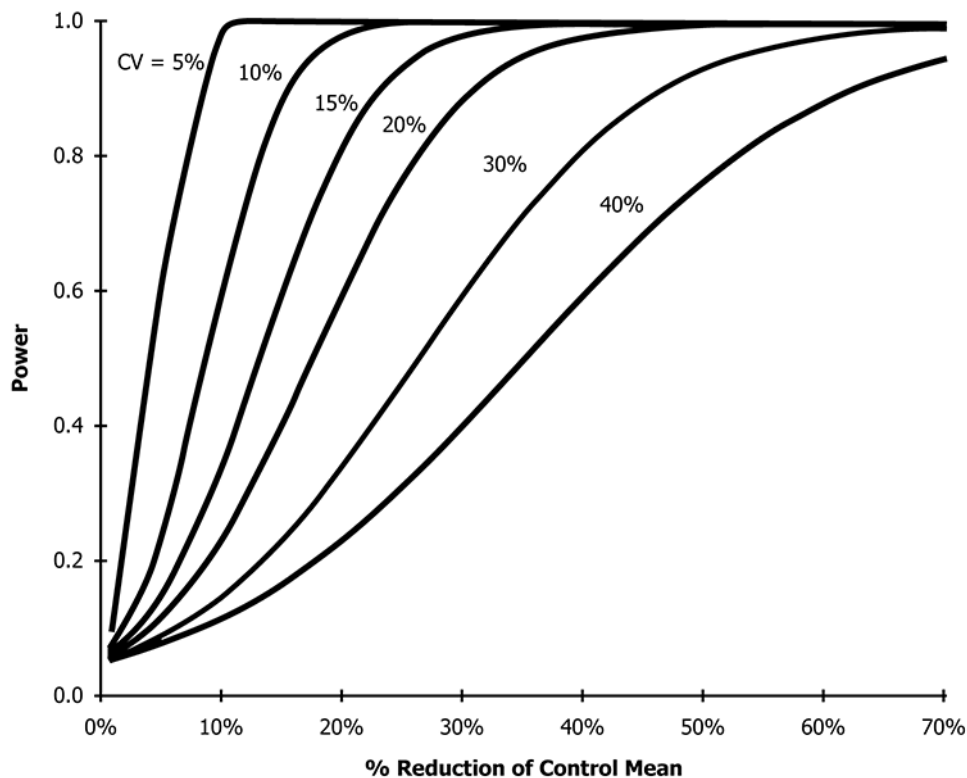
13.5.5.3 Error Types—Statistical tests of hypotheses can be designed to control for the chances of making incorrect decisions. In Fig. 5, alpha (α) represents the probability of making a Type I statistical error. A Type I statistical error in this testing situation results from the false conclusion that the treated sample is toxic relative to the control or reference sample. Beta (β) represents the probability of making a Type II statistical error, which is the erroneous conclusion that there are no differences between the mean responses in the treatment and control/reference sediments. Traditionally, values for α have ranged from 0.1 to 0.01 with 0.05 or 5 % used most commonly. Frequently, having chosen α , environmental researchers have ignored β and thus, the associated power of the test ($1-\beta$).

13.5.5.4 Variables Determining Error Rates—The critical components of the experimental design associated with the testing of hypotheses outlined above are: (1) the required MDD between the treatment and control responses or reference responses, (2) the variance among treatment and control or reference replicate experimental units, (3) the number of replicate units for the treatment and control or reference samples, (4) the number of animals exposed within a replicate exposure chamber, and (5) the selected probabilities of Type I (α) and Type II (β) errors.

13.5.5.5 Relationship Among MDD, Variability, and Power—To illustrate the principles of power analysis, Fig. 6 shows power curves for a two-sample t test with 8 replicates (α

= 0.05, one-tailed). The x -axis is the MDD is expressed as a percentage change from the mean control response. The y -axis is the power of the test, the likelihood that the text will correctly identify a difference between the treatment and the control or reference. Different curves are plotted representing different levels of inter-replicate variability, expressed as the coefficient of variation (CV; mean divided by standard deviation $\times 100$ %). For example, if the CV is 15 %, then a 20 % difference from control or reference can be detected with about 80 % probability (power ≈ 0.8). However, if the CV is higher at 30 %, then the same 20 % difference between treatment and control or reference will only be detectable about 20 % of the time (power ≈ 0.2). Clearly, inter-replicate variability is an important factor in test power, and is a reason test methods emphasize standardization and consistency as a means to minimize variability.

13.5.5.6 Influence of α and Replication on Power—Relaxing the α level of a statistical test increases the power of the test. In Fig. 6, α was set at 0.05. If α is relaxed to 0.1, then the chance of detecting a 20 % change from control or reference increases from about 80 % to 90 % for CV = 15 %, and from about 20 % to 30 % for CV = 30 %. Of course, increasing α also increases the chances of making a Type I error (false positive). Thus, the selection of α and its relationship to power must balance the costs associated with making Type I and II statistical errors. Reducing replication decreases power, and for that reason, numbers of replicates should only be reduced where there is reason to believe that power will not be unacceptably reduced.



NOTE 1—Alpha α = 0.05 represents the probability of making a Type I Error (false positive); Beta (β) represents the probability of making a Type II Error (false negative). Power is $1-\beta$, the chance of correctly determining that a difference exists.

FIG. 6 Power Curves for a Two-sample, One-tailed, t Test

13.5.5.7 Study-specific Calculation of Power—The power calculation shown in Fig. 6 is for a specific statistical test (two-sample *t* test) with $\alpha = 0.05$, and is presented only for illustrative purposes. The type of hypothesis test used (for example, number of comparisons) also influences power, and the permutations are too many to present effectively. Most common statistical software packages include routines that can perform study-specific power calculations for particular study designs, statistical tests, and anticipated variability.

13.5.6 Considerations for Replication when Using Regression Analysis:

13.5.6.1 Balancing the Number of Treatments Against Level of Replication—In the context of sediment toxicity testing, regression analysis is typically used for studies of spiked sediments, which allows assessment of organism response over a structured range of exposure concentrations. Strictly speaking, no replication at all is required for regression analysis, though that may not be advisable. Regression analysis of toxicity data is typically conducted using equations that have a sigmoid shape, with an asymptote to a constant (control or reference) response at low concentrations, transitioning to increasing effect at higher concentrations. The information needs for regressions of this type are treatments that define performance in the “no response” portion of the response curve, and treatments that show partial effect (for example, between 10 % and 90 % response), which define the slope of the response curve. As such, the number of treatments falling in the partial response range can be equally or more important than the level of replication for individual treatments. For example, if a study is to involve 40 total test chambers (not counting control or reference) to be tested over a fixed range in concentrations, the data are more likely to support robust regression analysis if the study uses 10 closely spaced concentrations with 4 replicates each, than if it used only 5 more widely spaced concentrations with 8 replicates each. That said, reducing replication of individual treatments will reduce statistical power if the analysis plan is to use both regression and hypothesis-based analyses.

13.5.6.2 Benefits of Replication in Regression Analysis—Although not technically required for regression analysis, replication of treatments does provide benefits. Replication sufficient to assess intra-replicate variability allows weighted regression techniques to be used, testing for deviations of the assumed model from the data, and the application of outlier tests to eliminate spurious data. Because the organism response corresponding to no effect is a key parameter for most regression approaches, increasing replication in the control or reference may be desirable to help ensure the “no effect” response is well defined. The impact of increasing replication for a given regression approach can be evaluated using data simulations, that is, generating multiple example data sets according to a defined true response curve and specified within treatment variability.

13.6 Comparison of Means Using Hypothesis-based Analysis:

13.6.1 Overview—Fig. 7 outlines a basic decision tree for analysis of survival, growth, or reproduction data using hypothesis testing. In the tests described herein, samples or

observations refer to replicates of treatments. Sample size *n* is the number of replicates (that is, exposure chambers) in an individual treatment, not the number of organisms in an exposure chamber. Overall sample size *N* is the combined total number of replicates in all treatments.

13.6.2 One-tailed Versus Two-tailed Tests—In formulating a statistical hypothesis, the alternative hypothesis can be one-sided (one-tailed test) or two-sided (two-tailed test). The null hypothesis (H_0) is always that the two values being analyzed are equal. A one-sided alternative hypothesis (H_a) is that there is a specified relationship between the two values (for example, one value is greater than the other) versus a two-sided alternative hypothesis (H_a) which is that the two values are simply different (that is, either larger or smaller). A one-tailed test is used when there is an *a priori* reason to test for a specific relationship between two means. This is often the case in toxicity tests, where the comparison is to determine whether performance (for example, survival, growth) is lower in the test sediment compared to control or reference (that is, the sample is toxic). A two-tailed test is used when differences in either direction are important to the study objectives. For the same number of replicates, one-tailed tests are more likely to detect statistically significant differences between treatments (for example, have a greater power) than are two-tailed tests.

13.6.3 Tests for Normality and Equality of Variances; Data Transformations—In general, parametric statistical analyses such as *t* tests and analysis of variance are appropriate only if: (1) there are independent, replicate experimental units for each treatment, (2) the observations within each treatment follow a normal distribution, and (3) variances for both treatments are equal or similar. The first assumption is an essential component of experimental design (13.5). The second and third assumptions can be tested using the data obtained from the experiment. Therefore, before conducting tests to compare treatments, tests for normality and equality of variances should be performed.

13.6.3.1 Testing for Normality—The most commonly used test for normality for small sample sizes ($N < 50$) is the Shapiro-Wilk’s test. This test determines if residuals are normally distributed. Residuals are the differences between individual observations and the treatment mean. Residuals, rather than raw observations, are tested because subtracting the treatment mean normalizes differences in absolute response among treatments. This scales the observations so that the mean of residuals for each treatment and over all treatments is zero. The Shapiro-Wilk’s test provides a test statistic *W*, which is compared to values of *W* expected from a normal distribution. *W* will generally vary between 0.3 and 1.0, with lower values indicating greater departure from normality. Because normality is desired, one looks for a high value of *W* with an associated probability greater than the pre-specified α level. Note that higher α levels are used when the number of replicates is small, or when the design is unbalanced, because these are the cases in which departures from normality have the greatest effects on *t* tests and other parametric comparisons (Table 9). If data fail the test for normality, even after transformation, nonparametric tests should be used for additional analyses (13.6.4.2 and 13.6.6, and 13.6.7).