J. N. Am. Benthol. Soc., 2003, 22(4):621–631

BRIDGES

BRIDGES is a recurring feature of J-NABS intended to provide a forum for the interchange of ideas and information between basic and applied researchers in benthic science. Articles in this series will focus on topical research areas and linkages between basic and applied aspects of research, monitoring, policy, and education. Readers with ideas for topics should contact Associate Editors, Nick Aumen and Marty Gurtz.

In this article, Stribling et al. discuss data quality issues to be considered when conducting taxonomic analyses for biological assessments. They differentiate between 2 broad areas of taxonomy—research and production taxonomic investigations—and consider how approaches to organism identification can vary between these 2 areas. The authors stress the importance of evaluating and communicating data quality, and that knowledge of quality assurance/quality control elements is essential before drawing conclusions from biological assessment results.

Nick Aumen, nick_aumen@nps.gov
Marty Gurtz, megurtz@usgs.gov
Co-editors

Determining the quality of taxonomic data

JAMES B. STRIBLING[1]
Tetra Tech, Inc., 10045 Red Run Blvd., Suite 110, Owings Mills, Maryland 21117-6102 USA

STEPHEN R. MOULTON II
US Geological Survey, National Water-Quality Assessment Program, 12201 Sunrise Valley Dr., 413 National Center, Reston, Virginia 20192 USA

GARY T. LESTER
EcoAnalysts, Inc., 105 East 2nd Street, Suite 1, Moscow, Idaho 83843 USA

A biological assessment protocol is a measurement system, a series of methods that functions to translate field samples to quantitatively based narrative assessments (Diamond et al. 1996, Barbour et al. 1999). Carter and Resh (2001) showed that there is widespread variability among monitoring programs in field and laboratory methods used in biological assessments. The components of any measurement system can contribute to the variability of the results (Taylor 1988, Warren-Hicks et al. 2000), so it is important for data users to understand the uncertainty associated with them. Therefore, it is necessary to evaluate individual components of the process, because each represents a potential error source (in this sense, we use *error* and *variability* interchangeably). The relative importance and acceptability of different error sources depends on specific objectives and data needs (Fig. 1); they typically are stated as measurement or data quality objectives (MQOs and DQOs, respectively) (Costanza et al. 1992, USEPA 2000).

Taxonomy is the theory and practice of classifying organisms (Mayr and Ashlock 1991). In assemblage-level biological assessments, taxonomy can add a degree of uncertainty to the result and can become even more critical if qualitative assessments are based on the presence or absence of particular taxa. We recognize 2 broad areas of taxonomic investigation—research and production (Table 1). We define research taxonomy (including biological systematics) as investigations leading to the description of new taxa or life stages, geographic range extensions, phylogenetic analyses, or documentation of autecological characteristics or morphological/anatomical structures. Research taxonomy typically is done in association with academic institutions such as universities and museums. We define production taxonomy as investigations (e.g., biological assessments) that process samples with the goal of producing a list of taxa and associated enumerations; the resulting data are then

[1] E-mail address: james.stribling@tetratech.com

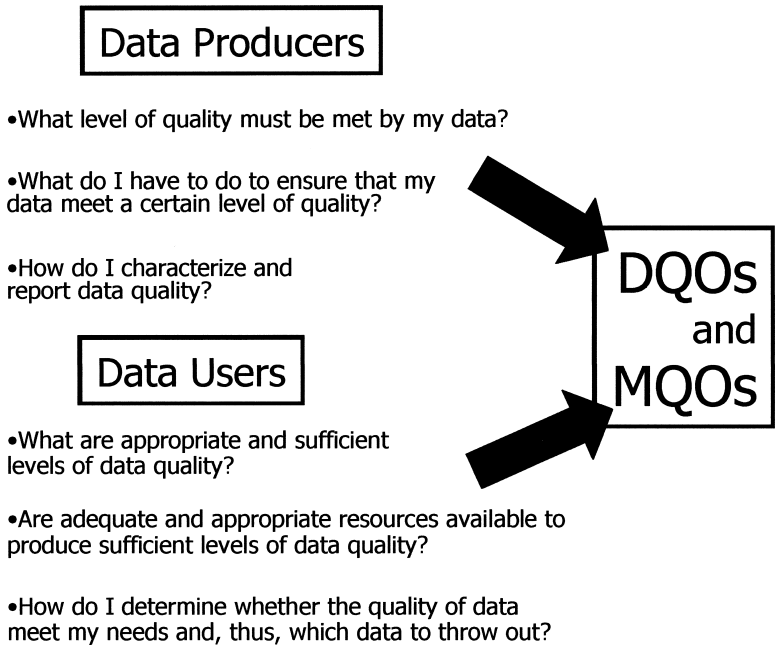

FIG. 1. The necessity and sufficiency of data quality are determined by users of the data; that data meet those needs is ensured by the data producers. These needs are communicated to data producers with data and measurement quality objectives (DQOs and MQOs).

used to address broader programmatic goals and objectives, i.e., production taxonomy is not the end in itself. Our use of the term production should not be confused with studies of biological productivity. Rather, it refers to business-like sample processing under the constraints of time, money, and deadlines, such as those described by Wilkie et al. (2003). Production taxonomy is commonly associated with government, academic, or environmental consulting laboratories that sponsor or support biological assessment studies. However, studies supported by production taxonomy have, on occasion, contributed to the discovery of new taxa and expanded distributional knowledge of organisms; the description and publishing of these findings is frequently done in collaboration with research taxonomists (e.g., Lester et al. 2002). Although many of the tools (e.g., experience and training, facilities, equipment, supplies, and technical literature) required for each kind of taxonomy are the same, how the tools are applied toward the different objectives can differ dramatically (Table 2).

This paper does not address different levels of taxonomic resolution necessary for detecting resource impairment (Ferraro et al. 1989, Ferraro and Cole 1992, Bailey et al. 2001, Lenat and Resh 2001); that is a different issue. We do stress, however, the importance of evaluating and communicating data quality (Costanza et al. 1992). Evaluation of taxonomic data quality involves both the communication of performance characteristics (e.g., accuracy and precision) and the documentation of operational factors (e.g., training, reference materials, protocols) under which the data are produced. We argue that knowledge of these quality assurance/quality control (QA/QC) elements are necessary for evaluation of taxonomic results used in biological assessments of water quality, or prior to drawing elaborate analytical conclusions using taxonomic data of unknown quality. Our discussion is focused on aquatic macroinvertebrate taxonomic data derived from discrete samples (Wilkie et al. 2003). The concept of evaluating taxonomic data quality should be applicable to other taxonomic groups such as algae and fishes, although the logistics for obtaining re-identifications will be substantially different. This approach has been used for botanical surveys (Scott and Hallam 2002). This paper is in-

TABLE 1. Distinctions between research and production taxonomy.

Topic	Type of taxonomy	
	Research	Production
Taxonomic focus	Work is generally focused in one taxon (e.g., Trichoptera, Hydropsychidae, or *Hydropsyche*)	Work includes examination of all taxa in a sample
Definition of a sample	Ranges from a single specimen to thousands of specimens collected at single station and date; sample generally collected to produce the highest abundance and (possibly) diversity of a single taxon; acquired quantitatively or qualitatively, depending on the question being addressed	The collection of organisms living in some aquatic substrate that are retained by the collection device or net; samples are unique with respect to station, date, time, habitat, and collection method; they are generally collected to produce a sample representative of the resident assemblage and may be acquired quantitatively or qualitatively, depending on the practical needs of the project
Condition of sample	Specimens most often in good condition; damaged specimens put aside until additional good specimens acquired; often focused on a specific life stage	Mixture of specimens in good and poor condition (damaged or early instar)
Origin of sample	Acquired from the field or from an existing collection (e.g., reference or voucher collections)	Acquired from the field
Goal	To address some question, problem, or need relative to the advancement of taxonomy	To produce a list of taxa, and usually, the associated counts in the sample
Time	Generally not a limiting factor to completing taxonomic identification	Extremely important, and often represents a major limiting factor to taxonomic resolution achieved
Taxonomic resolution	Almost always directed towards species-level distinctions	Spans entire nomenclatural hierarchy depending on numerous limiting factors (time, money, condition of specimens)
Type material	Comparison to type material very important for certain types of research (e.g., descriptions of new taxa)	Comparison to type material possible only with cooperation of a professional taxonomist

tended for all professional taxonomists, project managers in the process of designing and initiating biological monitoring programs, and natural-resource managers using taxonomic data for decision making.

Distinctions between Research and Production Taxonomy

Research taxonomy

The goal of research taxonomy is to advance our knowledge of the diversity and phylogenetic relationships of organisms. Research taxonomists provide the foundation of knowledge on which production taxonomists base their work. However, over the last several years there has been a noticeable decline in the number of research taxonomists despite an increased demand for taxonomic expertise (Cranston and Hillman 1992, New 1996). This increased demand is largely a result of an emphasis on the identification of aquatic organisms collected for biological assessments and the need to document the quality of these taxonomic data (Ellis and Cross 1981, Ellis 1988, Cranston and Hillman 1992, New 1996, Ranasinghe et al. 2003, Wilkie et al. 2003). If the availability of adequately trained research taxonomists continues to decline, then we can expect an eventual decrease in the quality of production taxonomy.

The work produced by research taxonomists is extensive and appears in several different forms of publication. Wiley (1981) listed 11

TABLE 2. The importance of operational factors and resource availability to elevating and maintaining the quality of taxonomic data (rating scale: 1 = low, 2 = moderate, 3 = high).

Operational factors and resource needs	Type of taxonomy	
	Research	Production
Training and experience	2	3
Technical literature	3	3
Internet access/online databases	3	3
Research collection	3	1
Reference collection	3	2
Voucher collection	1	3
Standard operating procedures	1	3
Standard data forms	1	3
Dissecting microscope	3	3
Compound/phase contrast microscope	3	3
Scanning electron microscope	3	1
Intralaboratory quality control	1	3
Interlaboratory quality control	1	3

kinds of taxonomic and systematics publications, including new taxa descriptions, revisions, classifications, keys, faunistic and floristic studies, atlases, catalogues, checklists, handbooks, taxonomic scholarship/rules of nomenclature, and phylogenetic analyses.

Production taxonomy

As distinct from research taxonomy, the primary goal of production taxonomy is to produce a list of names of specimens contained within a sample, and an enumeration of the number of individuals of each taxon in that sample. There are usually stringent time and funding restrictions for completion of the identifications, entry of the names and counts in a database, development of a voucher collection, and performance of all QC procedures. A foundation of high-quality research publications is crucial to production of high-quality taxonomic data. That information can be synthesized into a well-documented identification manual, set of dichotomous keys and illustrations, and relevant, taxon-specific data (such as functional feeding group and behavioral habit) (e.g., Usinger 1956, Brigham et al. 1982, Stehr 1987, 1991, Pennak 1989, Peckarsky et al. 1990, Merritt and Cummins 1996). These documents may help identify specimens only to family or genus level, and they are usually geographically or phylogenetically restricted. Additional technical literature is necessary to confirm that identifications are correct. Uncertainty in a final identification can be buffered by comparison to keys in ≥2 separate publications, particularly if different characters are used, or illustrations provided. The most current, up-to-date, and accepted (i.e., recognized by the community of benthic taxonomists to be valid, complete, and of high quality) publications should be used for identifications.

The quality of production taxonomy data sets requires that, in addition to key operational factors being maintained, the taxonomist is trained and experienced. Other important sources of error include specimens that are difficult-to-identify because they are early instar, damaged, or poorly mounted on slides. Unless the taxonomist is extremely familiar with the taxon, genus- or species-level identifications can be highly problematic. The less confident a taxonomist is with the group, the more likely s/he is to leave the identification at a less-refined level of hierarchical resolution, i.e., genus or family as opposed to species. In addition, time constraints often place severe restrictions on how much effort can be put into determining the identity of every specimen.

Documenting the Quality of Taxonomic Data

Performance characteristics

Performance characteristics that are used to document data quality include accuracy, precision, bias, and completeness (see definitions below). Although typically used in association with QA/QC activities, these 4 performance characteristics can be used to communicate the quality of an existing data set or express the level of data quality that would be acceptable to a data user (i.e., stating DQOs and MQOs).

There are 3 potential relationships between taxonomic accuracy and precision (Fig. 2). The 1st, and the goal of all taxonomic work, is that the data are both accurate and precise. The 2nd is that the data are precise but not accurate. This relationship represents a bias that might have resulted from the misinterpretation of a dichotomous key or morphological structure, or use of invalid nomenclature. Correction of this type of imprecision is easily managed by providing ad-

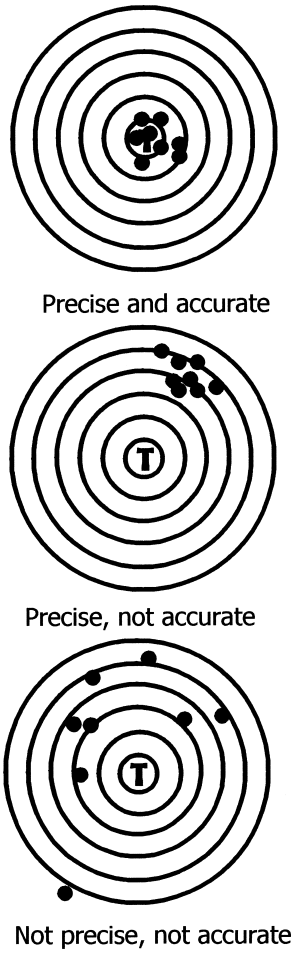

FIG. 2. Relationship between precision and accuracy. T represents the analytical truth (adapted from Keith 1991).

ditional training or acquisition of appropriate technical literature. The 3rd potential relationship represents an undesirable scenario where data are both inaccurate and imprecise. This type of relationship suggests the possibility of major operational problems with taxonomic competency and the availability of appropriate technical literature, reference specimens, and microscopes.

Accuracy.—Accuracy is defined as the nearness of a measurement to some analytical truth (Fig. 2), or true value (Taylor 1988, Clark and Whitfield 1994, Taylor and Kuyatt 1994, Random House 1996). The key to determining the accuracy of any measurement or measurement system is specification of the true value, which often can be the most difficult aspect of the process. In analytical chemistry, a reference standard of known concentration is the analytical truth (APHA 1995); thus, how well a laboratory method characterizes that standard is taken as the accuracy of the method/measurement.

For taxonomy, the analytical truth is: 1) the most currently accepted taxonomic literature, 2) a reference collection, preferably verified by appropriate taxonomic specialists, or 3) type material (e.g., holotype). Concomitantly, successful completion of a certification program can help demonstrate that a production taxonomist has the ability to recognize the analytical truth. Comparisons made with type material are practical only when there is potential for discovery of a new taxon or clarification of an important taxonomic problem, and is usually done in consultation with a research taxonomist.

Precision.—Precision is defined as the nearness of different measures of the same property (Taylor 1988, Taylor and Kuyatt 1994). More simply stated, it is a measure of repeatability. There are 2 approaches to characterizing precision: 1) using $\geq$2 methods to take a measurement of some property, and 2) using 1 method to measure that property repeatedly. In analytical chemistry, precision typically is determined through analysis and comparison of replicate or split samples (aliquots) (APHA 1995), with the differences usually represented as standard error, coefficient of variability, or relative % differences. The magnitude of those differences is a statement about the consistency or repeatability of the methods.

Taxonomic precision is evaluated by comparing the results of a randomly selected sample that is processed by 2 taxonomists or laboratories. The randomly selected samples represent a subset of the total collected for: 1) a project, 2) multiple projects within a sampling year, or 3) $\geq$1 projects over several sampling years, depending on the specific situation. The number of samples included in this subset can vary, but the rule-of-thumb for most projects is 10% of the total sample load. Precision can be quantified for both taxonomic identifications and enumerations. Upon receipt of re-identification results, 2 taxonomic lists are compared (Table 3).

Another aspect of taxonomic data quality is the final count of specimens for each taxon in the sample. Final specimen counts for samples depend on the taxonomic identifications, not the

TABLE 3. Example comparisons of re-identification results by 2 taxonomists showing counts of agreements. Target taxonomic level is based on program specifications.

Target taxonomic level	Identification	Taxonomist 1	Taxonomist 2	No. agreements
Genus	Baetidae		1	0
	Procloeon/Centroptilum	1		
Genus	*Argia*	1	2	1
	Coenagrionidae	1		
Genus	*Bratislavia*		2	2
	Bratislavia unidentata	2		
Genus	*Ceratopsyche morosa*	12		12
	Ceratopsyche bronta		12	
Genus	*Physa*		4	0
	Physidae	4		
Genus	*Dugesia tigrina*	1	25	1
	Cura foremanii	25		
Genus	*Glyptotendipes*	58	32	32
Species	*Polypedilum halterale*		9	0
	Polypedilum obtusum	9		
Genus	*Hexatoma*	4	4	4

rough counts obtained during the initial sorting activity. Precision of counts is determined by calculating % difference in enumeration (PDE) as follows:

$$\text{PDE} = \frac{|n1 - n2|}{n1 + n2} \times 100$$

where, $n1$ is the number of specimens counted in a sample by the 1st taxonomist or laboratory, and $n2$, the 2nd. The purpose of this calculation is to highlight those samples where counts might differ substantially and to focus attention on reason(s) for the miscounts.

Enumeration error can contribute to elevated uncertainty about data quality, but the extent to which it affects the ultimate use of the data can be minimal. However, documentation of PDE for a data set allows secondary users to evaluate its importance. One way to minimize sample enumeration error is to establish counting rules that will be used by the taxonomist. Examples include identifying and counting a specimen only if its head and ≥50% of its body are intact, or including a mollusk shell only if it is occupied by a specimen. Enumeration differences also will affect calculation of taxonomic precision.

Taxonomic results can be compared between 2 taxonomists or laboratories by counting the number of agreements, from which a % taxonomic disagreement (PTD) is calculated:

$$\text{PTD} = \left[1 - \left(\frac{comp_{pos}}{N}\right)\right] \times 100$$

where, $comp_{pos}$ is the number of agreements (positive comparisons), and N is the total number of specimens in the larger of the 2 counts. Agreements are, in part, contingent on the targeted level of identification, i.e., species, genus, family, or higher. For example, if genus is the target, and one taxonomist provides a name for a specimen at the species level, whereas the other leaves the name at genus level, it would be scored as an agreement. However, if one identification is at the genus level and the re-identification is at family, it would not be counted as an agreement (one identification met the target, the other did not). The lower the PTD value, the greater the overall taxonomic precision, indicating relative consistency in sample treatment. If disagreements affect a large number of specimens in either single or multiple samples throughout the entire data set, then those samples can be isolated and evaluated further for corrective re-identifications.

The MQO for the precision of a monitoring program or data set should be specified at a level acceptable to the data user. Uses of these val-

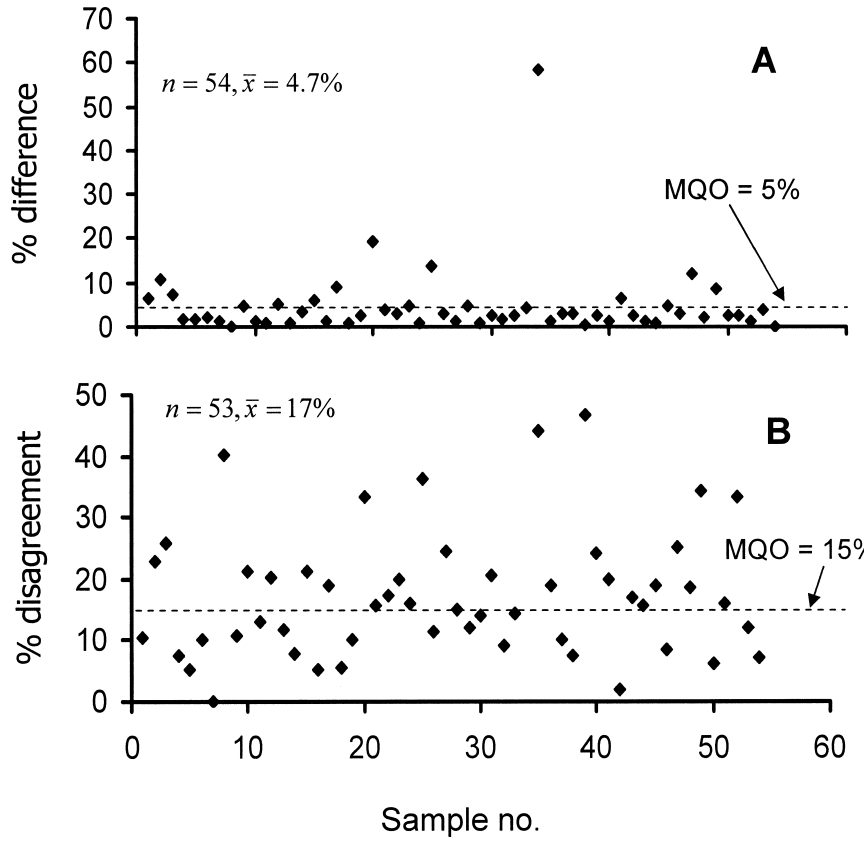

FIG. 3. Comparison of % difference in enumeration (A) and % taxonomic disagreement (B) for sample results that were reprocessed by 2 laboratories. The measurement quality objectives (MQO) are project-based.

ues can take several directions. For example, Fig. 3A shows that 12 samples exceeded the MQO for enumeration. The causes may have originated from over-cleared specimens mounted on slides, specimens lost during sample handling, or incorrect data recording or transcription. Specific examination of patterns within those 12 samples revealed that several problems were the result of Laboratory 1 failing to transfer slide-mounted specimens to Laboratory 2. A corrective action was to confirm that all slides, mostly Chironomidae and Oligochaeta, had been identified for all samples (not just those that were randomly selected). Another problem was that, on occasion, the 2nd taxonomist would find more specimens in the sample than the 1st taxonomist, which occurred because the 2nd taxonomist was unaware of the counting rules applied by the first taxonomist (the 2nd taxonomist counted all worm fragments as opposed to just those with heads). Because it was the 2nd taxonomist who counted fragments, this error had no effect on the interpretation of the non-QC samples. The need for greater clarity in how damaged specimens are counted also was identified. Overall, the data set had a PDE of 4.7% and demonstrated that an MQO of PDE = 5% was reasonable as a control limit.

In another example, 28 samples exceeded the 15% MQO for taxonomic precision (Fig. 3B). Patterns of disagreement within those samples revealed issues with misinterpretation of morphological features in the midge genera *Rheotanytarsus* and *Paratanytarsus*. This problem was revealed because in many of the duplicate results from the 28 samples, identical numbers of each of these genera were identified (i.e., one genus in the 1st set of results, and the other genus in the 2nd set of results). In many samples, the 1st taxonomist had identified 3 different genera of Amphipoda, and the 2nd taxonomist identified only *Hyalella* with a count equal to the total

number of individuals comprising the 3 genera identified by the 1st taxonomist. The QC analysis demonstrated an average PTD of 17%, indicating a need to examine those individual samples exceeding the MQO of 15% for suspect patterns. Resulting corrective actions included sending specimens of these 2 midge genera to a 3rd taxonomist for confirmation and consideration of the potential need to composite all amphipods to the order level.

Bias.—Error that exists in a consistently applied method is the systematic error, or bias, associated with that method or result (Smith et al. 1988, Clark and Whitfield 1994). Bias is defined as statistical or method error caused by systematically favoring some outcomes over others and can be characterized as the degree of departure from a true value. Bias in environmental measurement systems can be described as how far a result is from a known condition.

Taxonomic bias would exist if there were consistent misinterpretation of dichotomous keys or morphological features, poor processing of samples (e.g., poor slide-mounting technique), or inadequate optical equipment. If, for example, occasional problems with poor slide mounts are noted, the extent to which they influenced error in the taxonomic analysis should be evaluated. Bias could also be an issue if outdated or invalid dichotomous keys had been routinely used.

Completeness.—Completeness is a measure of the number of valid data points gathered relative to the number of planned data points (Smith et al. 1988). Completeness of biological assessments is typically viewed in terms of the number of samples that meet methodological specifications and requirements and can be used for analyses. We propose examining taxonomic completeness in a different manner.

Projects and protocols often specify targeted hierarchical levels to which specimens should be identified. (Standard operating procedures [SOPs] for some monitoring programs state that specimens are identified to the species level, but examination of data often reveals a high % of coarser-level identifications, which reflects that practical level can equate to the ability and background experience of the production taxonomist.) However, these targets frequently are missed because of specimen damage, early instar larvae, or poor slide mounting. In these instances, a note should be recorded on the identification bench sheet documenting the reasons why a target was not achieved (see Moulton et al. 2000). The frequency of taxonomic incompleteness arising because of specimens in poor condition decreases with increased training and experience of the taxonomist. Targeted taxonomic levels also can be missed because of taxonomic uncertainty resulting from unknown life stages or lack of knowledge about a particular taxon. Percent completeness can be calculated for each sample or as a mean for the data set overall.

Operational factors or conditions

The capacity or ability of a person or laboratory to produce high-quality taxonomic results is influenced by: 1) training and experience; 2) access to technical literature, reference and voucher collections, and taxonomic specialists; and 3) possession of appropriate and adequate optical equipment and laboratory facilities.

Training and experience.—Several upper-level undergraduate and graduate courses are taught/available that assist in development and maintenance of the ability to do taxonomic identifications competently. These courses are classroom or field-based and cover such diverse disciplines as taxonomy, morphology, anatomy, physiology, behavior, and ecology. Many of these courses will provide experience in: 1) proper use of dichotomous keys; 2) understanding and interpretation of morphology and morphological derivatives; 3) proper techniques for field sampling, specimen killing, preservation and preparation for identification, including dissections; and 4) development of reference and voucher collections. However, good taxonomic skills are not acquired quickly. Rather, they are developed over time by identifying many specimens, understanding and working with the taxonomic literature, and making comparisons to specimens in reference or voucher collections. Taxonomic certification by government agencies or professional technical organizations is another mechanism that documents the ability of a person to perform taxonomic identifications.

Literature, collections, and specialists.—Taxonomists rely on many resources to perform their work, in addition to his/her own training and experience. The types of literature that should be accessible are dichotomous keys, major text references, original descriptions, species checklists, and taxonomic revisions and reviews. It is

also helpful to consult taxonomic resources on the World Wide Web. For example, the Integrated Taxonomic Information System (ITIS; http://www.itis.usda.gov/) provides taxonomic hierarchy, authorship, verification status, nomenclatural validity, a voluntary taxonomic expert (name and contact number), geographic information, and links to other information regarding the taxon of interest. Misspelled taxonomic names in databases can result in calculation or analytical errors, and ITIS offers a straightforward mechanism of checking for these errors, especially for data-management specialists who are unfamiliar with the nomenclature.

Accessing specimen collections is extremely important in taxonomic work. Generally, 3 types of collections are recognized—research, reference, and voucher. A research collection can include the type specimens of different species, along with many other specimens. These collections generally are housed in major national, regional, or academic museums such as the US National Museum of Natural History, the Illinois Natural History Survey, and The Ohio State University Museum of Biological Diversity. Typically, only research taxonomists are allowed to examine specimens in these collections. A reference collection is a collection of identified specimens that is organized by major taxonomic grouping regardless of project or origin. When possible, reference specimens should exhibit the array of morphological characteristics, sexual differences, and life-cycle stages that define a particular taxon throughout its known distributional range. Reference collections typically have specimens that are, if available, in good condition and of sufficient developmental maturity to fully represent a particular taxon. The availability of a well-developed and well-curated reference collection provides powerful assistance in identification of unknown specimens and in verifying identifications. At a minimum, reference specimens should be verified internally as part of a laboratory QA program; ideally, these specimens also should be verified by an outside specialist. A voucher collection consists of identified specimens organized by sample and project. Voucher collections are important because they facilitate the re-identification of a discrete sample. However, since voucher specimens are sometimes deposited in a reference collection, a special note should be added to the bench data sheet indicating the final disposition of the specimens. Voucher collections are either discarded after some specified length of time (usually determined by a project) or archived indefinitely at the processing laboratory or some other location.

Even after using appropriate literature resources or making comparisons to reference specimens, it might be necessary to consult a taxonomic specialist for guidance on a difficult identification or taxonomic issue. These specialists usually are quite helpful and welcome inquiries to verify specimens or provide information. However, it is customary to obtain permission of the specialist before sending specimens to them for examination. Consultation fees might apply under certain circumstances, depending on the amount of work involved. If the specialist finds some interesting details about the specimens (e.g., a new distributional record or taxon), then they might ask to retain one or several specimens for other work in which they are involved.

Laboratory conditions and methods.—Appropriate optical equipment is necessary for performing identifications. However, the degree of magnification necessary will vary according to the taxonomic group, specimen condition, and the required level of specificity. Thus, dissecting (with variable reflective base) and compound microscopes with a range of ocular lenses will afford the greatest flexibility to a laboratory doing taxonomic analyses on large, multitaxon samples. Moulton et al. (2000) provided a list of supplies often necessary to do invertebrate taxonomy. Taxonomic laboratories should have a QA program in place to ensure high-quality data, including reviewed and up-to-date SOPs that detail acceptable technical literature and nomenclature (or the means to determine acceptability), when and how dissections are performed, and a list of taxa included in the reference collection.

Performance Characteristics and the Acceptability of Taxonomic Results

When comparing taxonomic results across different sets of samples, a key source of error is the use of different taxonomists. This problem can arise because individuals have been reassigned, the volume of work is too large for one person, labor turnover has occurred during the period of interest, or an individual has limited training and experience. How much these changes affect the overall quality of the taxonomy is usually unknown. We suggest that taxonomic reports

should include documentation regarding performance characteristics, the quality of the taxonomic data, and acceptance criteria. Moulton et al. (2000) present an approach for continuous QC oversight and corrective actions.

The users of taxonomic data (data analysts, natural resource assessors and managers, and secondary users) are ultimately responsible for deciding whether data are adequate for their needs (Fig. 1), and whether poor data are better than no data at all (Scott and Hallam 2002). The user should be able to understand data quality in the context of performance characteristics. Otherwise, they must be satisfied with assurances from the data producer (Costanza et al. 1992), a situation that reduces the objectivity and defensibility of biological assessment data, results, and interpretation. Production taxonomists must be equipped with necessary and adequate facilities, equipment, supplies, and literature; they also must have appropriate training and experience. However, the fact that facilities and conditions are available to someone performing production taxonomy does not ensure that accurate nomenclature is consistently applied to every specimen in every sample. Not knowing the extent to which it *is* consistently applied increases the risk of incorrect decisions. We suggest that data users need to exercise caution when basing biological assessments on taxonomic data sets that do not specifically present documentation of at least some performance characteristics.

Acknowledgements

We thank Marty Gurtz, John Morse, Mark Wetzel, Bern Sweeney, Scott Grotheer, Mike Winnell, Joe Flotemersch, Karen Blocksom, Kristen Pavlik, Matt Hicks, Dave Loch, Jerry Diamond, and Mike Barbour for their helpful comments on earlier drafts of the manuscript. USGS colleague reviews were provided by Daren Carlisle, Daniel Hippe, and Barry Poulton. Nick Aumen, Pamela Silver, Dave Rosenberg, and 2 anonymous reviewers also provided comments that helped improve the manuscript.

Literature Cited

APHA (AMERICAN PUBLIC HEALTH ASSOCIATION). 1995. Standard methods for the examination of water and wastewater. 19th edition. American Public Health Association, Washington, DC.

BAILEY, R. C., R. H. NORRIS, AND T. B. REYNOLDSON. 2001. Taxonomic resolution of benthic macroinvertebrate communities. Journal of the North American Benthological Society 20:280–286.

BARBOUR, M. T., J. GERRITSEN, B. D. SNYDER, AND J. B. STRIBLING. 1999. Revision to the rapid bioassessment protocols for streams and wadeable rivers: periphyton, benthic macroinvertebrates and fish. EPA 841-D-97-002. Office of Water, US Environmental Protection Agency, Washington, DC.

BRIGHAM, A. R., W. U. BRIGHAM, AND A. GNILKA (EDITORS). 1982. Aquatic insects and oligochaetes of North and South Carolina. Midwest Aquatic Enterprises, Mahomet, Illinois.

CARTER, J. L., AND V. H. RESH. 2001. After site selection and before data analysis: sampling, sorting, and laboratory procedures used in stream benthic macroinvertebrate monitoring programs by USA state agencies. Journal of the North American Benthological Society 20:658–682.

CLARK, M. J. R., AND P. H. WHITFIELD. 1994. Conflicting perspectives about detection limits and about the censoring of environmental data. Water Resources Bulletin 30:1063–1079.

COSTANZA, R., S. O. FUNTOWICZ, AND J. R. RAVETZ. 1992. Assessing and communicating data quality in policy-relevant research. Environmental Management 16:121–131.

CRANSTON, P., AND T. HILLMAN. 1992. Rapid assessment of biodiversity using 'Biological Diversity Technicians'. Australian Biologist 5:144–154.

DIAMOND, J. M., M. T. BARBOUR, AND J. B. STRIBLING. 1996. Characterizing and comparing bioassessment approaches and their results: a perspective. Journal of the North American Benthological Society 15:713–727.

ELLIS, D. V. 1988. Quality control of biological surveys. Marine Pollution Bulletin 19:506–512.

ELLIS, D. V., AND S. F. CROSS. 1981. A protocol for interlaboratory calibrations of biological species identifications (ring tests). Water Research 15:1107–1108.

FERRARO, S. P., AND F. A. COLE. 1992. Taxonomic level sufficient for assessing a moderate impact on macrobenthic communities in Puget Sound, Washington, USA. Canadian Journal of Fisheries and Aquatic Sciences 49:1184–1188.

FERRARO, S. P., F. A. COLE, W. A. DEBEN, AND R. C. SWARTZ. 1989. Power-cost efficiency of eight macrobenthic sampling schemes in Puget Sound, Washington, USA. Canadian Journal of Fisheries and Aquatic Sciences 46:2157–2165.

KEITH, L. H. 1991. Principles of environmental sampling and analysis: a practical guide. Lewis Publishers, Chelsea, Michigan.

LENAT, D. R., AND V. H. RESH. 2001. Taxonomy and stream ecology—The benefits of genus- and species-level identifications. Journal of the North American Benthological Society 20:287–298.

LESTER, G. T., W. P. MCCAFFERTY, AND M. R. EDMONDSON. 2002. New mayfly (Ephemeroptera) records from Idaho. Entomological News 113:131–136.

MAYR, E., AND P. D. ASHLOCK. 1991. Principles of systematic zoology. McGraw-Hill, New York.

MERRITT, R. W., AND K. W. CUMMINS (EDITORS). 1996. An introduction to the aquatic insects of North America. 3rd edition. Kendall-Hunt Publishing Company, Dubuque, Iowa.

MOULTON, S. R., J. L. CARTER, S. A. GROTHEER, T. F. CUFFNEY, AND T. M. SHORT. 2000. Methods of analysis by the U.S. Geological Survey National Water Quality Laboratory—Processing, taxonomy, and quality control of benthic macroinvertebrate samples. US Geological Survey Open-File Report 00–212. National Water Quality Laboratory, US Geological Survey, Denver, Colorado.

NEW, T. R. 1996. Taxonomic focus and quality control in insect surveys for biodiversity conservation. Australian Journal of Entomology 35:97–106.

PECKARSKY, B. L., P. R. FRAISSINET, M. A. PENTON, AND D. J. CONKLIN. 1990. Freshwater invertebrates of northeastern United States. Cornell University Press, Ithaca, New York.

PENNAK, R. W. 1989. Fresh-water invertebrates of the United States. 3rd edition. John Wiley and Sons, New York.

RANASINGHE, J. A., D. E. MONTAGNE, S. B. WEISBERG, M. BERGEN, AND R. G. VELRADE. 2003. Variability in the identification and enumeration of marine benthic invertebrate samples and its effect on benthic assessment measures. Environmental Monitoring and Assessment 81:199–206.

RANDOM HOUSE. 1996. Webster's encyclopedic unabridged dictionary of the English language. Gramercy Books, New York.

SCOTT, W. A., AND C. J. HALLAM. 2002. Assessing species misidentification rates through quality assurance of vegetation monitoring. Plant Ecology 165: 101–115.

SMITH, F., S. KULKARNI, L. E. MYERS, AND M. J. MESSNER. 1988. Evaluating and presenting quality assurance data. Pages 157–168 *in* L. H. Keith (editor). Principles of environmental sampling. ISBN 0-8412-1173-6. American Chemical Society, Salem, Massachusetts.

STEHR, F. W. (EDITOR). 1987. Immature insects. Volume 1. Kendall-Hunt Publishing Company, Dubuque, Iowa.

STEHR, F. W. (EDITOR). 1991. Immature insects. Volume 2. Kendall-Hunt Publishing Company, Dubuque, Iowa.

TAYLOR, B. N., AND C. E. KUYATT. 1994. Guidelines for evaluating and expressing the uncertainty of NIST measurement results. NIST Technical Note 1297. National Institute of Standards and Technology, US Department of Commerce, Washington, DC.

TAYLOR, J. K. 1988. Defining the accuracy, precision, and confidence limits of sample data. Pages 102–107 *in* L. H. Keith (editor). Principles of environmental sampling. ISBN 0-8412-1173-6. American Chemical Society, Salem, Massachusetts.

USINGER, R. L. 1956. Aquatic insects of California. University of California Press, Berkeley.

USEPA (US ENVIRONMENTAL PROTECTION AGENCY). 2000. Guidance for the data quality objective process (EPA QA/G-4). EPA/600/R-96/005. Quality Assurance Management Staff, Office of Environmental Information, US Environmental Protection Agency, Washington, DC. (Available from: http://www.epa.gov/quality/qs-docs/g4-final.pdf)

WARREN-HICKS, W. J., B. R. PARKHURST, D. R. J. MOORE, R. S. TEED, R. B. BAIRD, R. BERGER, D. L. DENTON, AND J. J. PLETL. 2000. Assessment of whole effluent toxicity test variability: partitioning sources of variability. Environmental Toxicology and Chemistry 19:94–104.

WILEY, E. O. 1981. Phylogenetics: the theory and practice of phylogenetic systematics. John Wiley and Sons, New York.

WILKIE, L., G. CASSIS, AND M. GRAY. 2003. A quality control protocol for terrestrial invertebrate biodiversity assessment. Biodiversity and Conservation 12:121–146.

ABOUT THE AUTHORS

James Stribling is a biologist and Associate Director for Tetra Tech's Center for Ecological Sciences. He does biological monitoring and assessment, develops approaches for documenting field and laboratory data-quality characteristics and method comparability, and coordinates development of monitoring designs with programmatic goals and objectives. Steve Moulton is the Biological Operations Manager for the National Water-Quality Assessment (NAWQA) Program of the US Geological Survey. He coordinates support for laboratory and field operations, training, and quality assurance/quality control activities related to taxonomic data collected by the NAWQA program. Gary Lester is President and CEO of EcoAnalysts, Inc., an aquatic resource-consulting firm. In addition to his administrative responsibilities, he does Chironomidae taxonomy, biological assessments, and develops and oversees laboratory sorting and identification quality control procedures.