Application of the Index of Biotic Integrity to Evaluate Water Resource Integrity in Freshwater Ecosystems

Thomas P. Simon and John Lyons

1.0 INTRODUCTION

The biotic and abiotic: processes involved in the environmental degradation of freshwater ecosystems are often complex, and their combined effects are not easily measured. Many human activities across the landscape contribute to degradation, including discharge of domestic, agricultural, and industrial effluents; cultural eutrophication; acidification; erosion and sedimentation following human modification of watersheds; straightening, deepening, and clearing of stream channels; drainage of wetlands and impoundments of streams; flow alterations caused by dam operation and water diversions; overharvest of biota; and introduction of nonnative species. Estimating ecosystem health and integrity may be the best way to assess the total effects of these activities on aquatic environments (Karr 1991).

Fish communities are a highly visible and sensitive component of freshwater ecosystems, and have several attributes that make them useful indicators of biological integrity and ecosystem health (Table 1). Fish communities respond predictably to changes in both abiotic factors, such as habitat and water quality (e.g., Karr 198 1; Hughes 1985; Karr et al. 1986; Leonard and Orth 1986; Scott et al. 1986; Berkman and Rabeni 1987; Steedman 1988; Simon 1990), and biotic factors, such as human exploitation and species additions (e.g., Hartman 1972; Colby et al. 1987; Rincon et al. 1990; Ross 1991). Many studies have identified the specific responses of fish communities to particular types of degradation (e.g., Forbes and Richardson 1913; Hubbs 1933; Millet et al. 1966; Gammon 1976; Karr et al. 1985a, b; Hughes et al. 1990; ORSANCO 1991; Hite et al. 1992).

Beginning around 1900 and accelerating greatly in the last 20 years, fish community characteristics have been used to measure relative ecosystem health (Fausch et al. 1990). Recent advances are attributed to the development of integrative ecological indices that directly relate fish communities to other biotic and abiotic components of the ecosystem (Karr 198 1; Karr et al. 1986; Ohio EPA 1987a,b; Plafkin et al. 1989; Barbour et al., Chapter 6), the delineation of ecoregions that allow explicit consideration of natural differences among fish communities from different geographic areas (Hughes et al. 1986, 1987; Omernik 1987; Hughes and Larsen 1988; Omernik, Chapter 5), and the recognition of the importance of cumulative effects of degradation at the landscape scale (Hughes, Chapter 4; Larsen et al., Chapter 18).

In the United States, biological criteria or standards based on fish communities have been formulated by several state and federal agencies to assess and protect freshwater ecosystem health (Ohio EPA 1987a,b; Karr 1991; Southerland and Stribling Chapter 7). A variety of quantitative indices can define specific biocriteria including indicator species or guilds; species richness, diversity, and similarity indices; the Index of Well-Being; multivariate ordination and classification; and the, Index of Biotic Integrity (reviewed by Fausch et aL 1990). Of these, the most commonly used and arguably the most, effective has been the Index of Biotic Integrity or IBI.

In this paper, we review the evolution and use of the IBI, with particular emphasis on recent applications. Karr et al. (1986), Miller et al. (1987), and Fausch et al. (1990) have reviewed the early development of the IBI, but since these publications, several major "new" versions have been developed. Many new versions are not documented in the primary, peer-reviewed literature, and thus are unknown to most water resource professionals, although these versions are being used to make important management decisions in many states. Our goal is to describe the different ways in which the original IBI has. been modified for use in different geographic regions and in different types of freshwater ecosystems.

With many different versions now in existence, the IBI is best thought of as a family of related indices rather than a single index. The IBI includes attributes of the biota that range from individual health to population, community, and ecosystem levels. We define the IBI broadly, as any index that is based on the sum or ratings for several different measures, termed metrics, of fish structure and/or function, with the rating for each metric based on quantitative expectations of what comprises high biotic integrity. For some metrics, expectations will vary depending on ecosystem size and location. The IBI is not a community analysis but it is an analysis of several hierarchical levels of biology that uses a sample of the assemblage.

Table 1. Attributes of Fishes that Make Them Desirable Components of Biological Assessments and Monitoring Programs

Goal/quality	Attributes
Accurate assessment of environmental health	* Fish populations and individuals generally remain in the same area during summer seasons * Communities are persistent and recover rapidly from natural disturbances * Comparable results can be expected from an unperturbed site at various times * Fish have large ranges and are less affected by natural microhabitat differences than smaller organisms. This makes fish extremely useful for assessing regional and macrohabitat differences * Most fish species have long life spans (2-10+ years) and can reflect both, long-term and current water resource quality * Fish continually inhabit the receiving water and integrate the chemical, physical, and biological histories of the waters * Fish represent a broad spectrum of community tolerances from very sensitive to highly tolerant and respond to chemical, physical, and biological degradation in characteristic response patterns
Visibility	* Fish are highly visible and valuable components of the aquatic community to the public * Aquatic life uses and regulatory language are generally characterized in terms of fish (i.e., fishable and swimmable goal of the Clean Water Act)
Ease of Use and Interpretation	* The sampling frequency for trend assessment is less than for short-lived organisms * Taxonomy of fishes is well established, enabling professional biologists the ability to reduce laboratory time by identifying many specimens in the field * Distribution, life histories, and tolerances to environmental stresses of many species of North American fish are documented in the literature

Modified from Plafkin et al. 1989; Simon 1991.

2.0 A BRIEF HISTORY OF THE INDEX OF BIOTIC INTEGRITY

The Index of Biotic Integrity (IBI) was first developed by Dr. James Karr for use in small warmwater streams (ie., too warm to support salmonids) in central Illinois and Indiana (Karr 1981). The original version had 12 metrics that reflected fish species richness and composition, number and abundance of indicator species, trophic organization and function, reproductive behavior, fish abundance, and condition of individual fish (Table 2). Each metric received a score of five points if it had a value similar to that expected for a fish community characteristic of a system with little human influence, a score of one point if it had a value similar to that expected for a fish community that departs significantly from the reference condition, and a score of three points if it had an intermediate value. Sites with high biotic integrity had relatively high numbers of total species, sucker (Catostomidae) species, darter *(Crystallaria, Ammocrypta, Etheostoma, and Percina;* Percidae), sunfish (Centrarchidae excluding *Micropterus)* species, and intol- erant species; high relative abundance of top carnivores and insectivorous cyprinid species; high overall fish abundance; and low relative abundance of the tolerant green sunfish *(Lepomis cyanellus),* omnivores, hybrids, and fish with diseases or deformities. Expectations for species richness metrics increased with increasing stream order, and were derived from an empirical relationship between stream size and maximum number of species present, termed the maximum species richness (MSR) line (Fausch et al. 1984). The total IBI score was the sum of the 12 metric scores and ranged from 60 (best) to 12 (worst) (some authors have reduced the lowest score to zero, e.g., Simon 1991).

The original version of the IBI quickly became popular, and was used by many investigators to assess warmwater streams throughout the central United States (e.g., Berkman et al. 1986; Gorman 1987a, b; Bickers et al. 1988; Hite and Bertrand 1989; Simon 1990; Hite et al. 1992; Osborne et al. 1992). Karr and colleagues explored the sampling properties and effectiveness of the original version in several different regions and different types of streams (Fausch et al. 1984; Karr et al. 1986).

As the IBI became more widely used, different versions were developed for different regions and different ecosystems. These new versions had a multimetric structure, but differed from the original version in the number, identity, and scoring of metrics (Table 2). New versions developed for streams and rivers in the central United States generally retained most of the metrics used in the original IBI, modifying only those few that proved insensitive to environmental degradation in a particular geographic area or type of stream (e.g., Whittier et al. 1987; Saylor et al. 1988; Crumby et al. 1990; Rankin and Yoder 1990; Saylor and AhIstedt 1990; Simon 1992; Hoefs and Boyle 1992; Lyons 1992). However, new versions developed for streams and rivers in France, Canada, and the eastern and western United States tended to have a very different set of metrics (e.g., Moyle et al. 1986; Schrader 1986; Langdon 1989; Steedman 1988; Bramblett and Fausch 1991; Oberdorff and Hughes 1992; Goldstein et al. 1994), reflecting, the substantial differences in fish faunas between these regions and the central United States. Similarly, the metrics used in IBI versions developed for other types of ecosystems, such as estuaries, impoundments, and natural lakes, usually bore only a limited resemblance to those of the original version (Thompson and Fitzhugh 1986; Thoma 1990; Dionne and Karr 1992; Hughes et al. 1992; Dycus and Meinert 1993; Larsen and Christie 1993; but see Greenfield and Rogner 1984 for an exception) yet retain the ecological structure of the original IBI metrics.

Table 2. List of Original Index of Biotic Integrity Metrics (Capital Letters) Proposed by Karr (1981) for Streams in the Central United States, Followed by Modifications Proposed by Subsequent Authors for Streams in Other Regions or for Different Streams and River Types

Species Richness and Composition Metrics

1. Total Number of Fish Species (1-7, 11, 13-15, 18, 19, 22)
   A. Number of native fish species (8-10, 12, 16, 17, 20, 22)
   B. Number of fish species, excluding Salmonidae (13)
   C. Number of amphibian species (3, 13)

2. Number of Catostomidae Species (1, 4-6, 8, 9, 12, 15, 17, 19, 20)
   A. Percent of individuals that are Catostomidae (9)
   B. Percent of individuals that are round-bodied Catostomidae: *Cycleptus*, *Hypentelium,* *Minytrema,* and *Moxostoma* (9, 17, 19)
   C. Number of Catostomidae and Ictaluridae species (10)
   D. Number of Catostomidae and Cyprinidae species (17)
   E. Number of benthic insectivorous species (7, 11, 17, 22)
   F. Number of laterally compressed minnow species (21)
   G. Number of minnow species (4, 6, 9, 14, 15, 17, 22)
   H. This metric deleted from IBI (2, 3, 13, 14)
3. Number of Darter Species (Percidae genera: *Crystallaria*, *Etheostoma, P*ercina, *Ammocrypta*) (1, 2, 4, 5, 8, 9, 12, 15-17, 19, 22)
   A. Number of darter and Cottidae species (9, 10)
   B. Number of darter, Cottidae, and *Noturus* (Ictaluridae) species (15, 16, 19)
   C. Number of darter, Cottidae, and round-bodied Catostomidae species (17)
   D. Number of Cottidae species (6, 13)
   E. Abundance of Cottidae individuals (3)
   F. Number of benthic species (11, 18)
   G. Percent of individuals that are native benthic species (11)(same as #2)
   H. Number of benthic insectivorous species (7)
   I. Number of darter species, excluding “tolerant darter species” (headwater sites) (21)
   J. Percent cyprinids with subterminal mouths (22) K. This metric deleted from IBI (14)
4. Number of Sunfish Species (Centrarchidae excluding *Micropterus*)(1, 4, 5, 14-17, 20)
   A. Number of native sunfish species (9, 12)
   B. Number of sunfish and Salmonidae species (10)
   C. Number of sunfish species and *Perca flavescens* (Percidae) (16)
   D. Number of headwater (restricted to small streams) species (9, 15, 22)
   E. Number of water column (non-benthic) species (7, 17, 18)
   F. Number of water column cyprinid species (17)
   G. Number of sunfish species including *Micropterus* (19, 22)
   H. This metric deleted from IBI (11)

Indicator Species Metrics

5. Number of Intolerant or Sensitive Species (1-4, 6-8, 10-12, 14, 15, 17-20, 22)
   A. Number of Salmonidae species (3,15)
   B. Percent of individuals that are Salmonidae (11)
   C. Juvenile Salmonidae presence or abundance (3, 18)
   D. Large (>1 5-20 cm) or adult Salmonidae presence or abundance (3, 6, 18)
   E. Abundance or biomass of all sizes of Salmonidae (3, 13)
   F. Mean length or weight of Salmonidae (13)
   G. Percent of individuals that are anadromous *Oncorhynchus mykiss* (Salmonidae) older than age 1 (3)
   H. Presence of *Salvelinus fontinalis* (Salmonidae) (10)
   I. Presence of juvenile or large *Esox lucius* (Esocidae) (18)
   J. Number of Large River (restricted to great rivers) species (19, 22)
   K. Percent of species that are native species (3)
   L. Percent of individuals that are native species (3)
   M. This metric deleted from IBI (2, 14)

6. Percent of Individuals that Are *Lepomis cyanellus (Centrarchidae)* (1, 17)
   A. Percent of individuals that are *Lepomis megalotis* (Centrarchidae) (5)
   B. Percent of individuals that are *Cyprinus carpio* (Cyprinidae) (6)
   C. Percent of individuals that are *Semotilus atromaculatus* (Cyprinidae)(2)
   D. Percent of individuals that are *Rutilus rutilus* (Cyprinidae) (18)
   E. Percent of individuals that are *Rhinichthys species* (Cyprinidae) (10)
   F. Percent of individuals that are *Catostomus commersoni* (Catostomidae) (4, 7, 11)
   G. Percent of individuals that are tolerant species (8, 9, 12, 14-17, 19, 22)
   H. Percent of individuals that are “pioneering” species (9, 15, 22)
   I. Percent of individuals that are introduced species (4, 6, 12, 14)
   J. Number of introduced species (12, 13)
   K. Evenness (22)
   L. This metric deleted from IBI (3)

Table 2 (continued). List of Original Index of Biotic Integrity Metrics (Capital Letters) Proposed by Karr (1981) for Streams in the Central United States, Followed by Modifications Proposed by Subsequent Authors for Streams in Other Regions or for Different Streams and River Types

Trophic Function Metrics

7. Percent of Individuals that Are Omnivores (1-4, 6-8, 10, 12, 14-20, 22)
   - A. Percent of individuals that are omnivorous Cyprinidae species (10)
   - B. Percent of individuals that are *Luxilus cornutus* or *Cyprinella spiloptera* (Cyprinidae), facultative omnivores (9)
   - C. Percent of individuals that are generalized feeders that eat a wide range of animal material but limited plant material (2, 9, 11)
   - D. Percent biomass of omnivores (22)
   - E. This metric deleted from IBI (3, 13)

8. Percent of Individuals that Are Insectivorous Cyprinidae (1, 17)
   - A. Percent of individuals that are insectivores/invertivores (5-7, 9, 12, 14-19)
   - B. Percent of individuals that are specialized insectivores (2, 4, 20)
   - C. Percent of individuals that are specialized insectivorous minnows and darters (8)
   - D. Percent biomass of insectivorous cyprinids (22)
   - E. This metric deleted from IBI (3, 10, 13)

9. Percent of Individuals that Are Top Carnivores or Piscivores (1, 5, 7-9, 11, 12, 15-20)
   - A. Percent of individuals that are large (> 20 cm) piscivores (10)
   - B. Percent biomass of top carnivores (22)
   - C. This metric deleted from IBI (2-4, 6, 13-15)

Reproductive Function Metrics

10. Percent of Individuals that Are Hybrids (1, 7, 8, 13)
   - A. Percent of individuals that are simple lithophilous species: spawn on gravel, no nest, no parental care (9, 15-17, 19, 20, 22)
   - B. Percent of individuals that are gravel spawners (18)
   - C. Ratio of broadcast spawning to nest building cyprinids (22)
   - D. This metric deleted from IBI (2-6, 10-12, 14)

Abundance and Condition Metrics

11. Abundance or Catch per Effort of Fish (1 -8, 10-11, 14-15, 18, 19, 22)
   - A. Catch per effort of fish, excluding tolerant species (9, 16, 20)
   - B. Biomass of fish (6, 13, 22)
   - C. Biomass of amphibians (13)
   - D. Density of macroinvertebrates (13)
   - E. This metric deleted from IBI (17)

12. Percent of Individuals that are Diseased, Deformed, or Have Eroded Fins, Lesions, or Tumors (1-7, 9, 11, 12, 14-16, 18, 19, 20, 22)
   - A. Percent of individuals with heavy infestation of cysts of the parasite *Neascus* (10)
   - B. This metric deleted from the IBI (13, 17)

Note: The numbers in parentheses correspond to the following references.

1. Karr (1981); Fausch et al. (1984); Karr et al. (1985a,b); Karr et al. (1986); Angermeier and Karr (1986); Berkman et al. (1986); Karr et al. (1987); Hite and Bertrand (1989); Angermeier and Schlosser (1988); Hite et al. (1992); Osbome et al. (1992).
2. Leonard and Orth (1986).
3. Moyle et al. (1986).
4. Schrader (1986).
5. Foster (1987).
6. Hughes and Gammon (1987).
7. Miller et al. (1988).
8. Saylor and Scott (1987); Saylor et al. (1988); Saylor and Ahlstedt (1990).
9. Ohio EPA (1987a,b).
10. Steedman (1988).
11. Langdon (1989).
12. Crumby et al. (1990).
13. Fisher (1990).
14. Bramblet and Fausch (1991).
15. Simon (1991).
16. Lyons (1992).
17. Hoefs and Boyle (1992).
18. Oberdorff and Hughes (1992).
19. Simon (1992).
90. Bailey et al. (1993).
21. Gatz and Harig (1993).
22. Goldstein et al. (1994).

3.0 CRITICAL FLOW VALUES AND BIOLOGICAL INTEGRITY

Water quality standards contain rules that define minimum stream flows above which chemical and narrative criteria must be met. This is most commonly the seven-day average flow that has a probability of recurring once every ten years (i.e., $Q_{7,10}$ flow). Other low-flow values can be used (95% duration flow, to the annual hydrograph for a given $Q_{30,10}$ flow) as well and these can approximate the $Q_{7,10}$ relative to the annual hydrographic for a given stream or river. Because the customary use of chemical and narrative criteria is essentially based on a steady-state, dilution-oriented process, a design "critical" flow is necessary. This has been widely accepted and essentially unquestioned practice in surface water quality regulation for many years. It is an inherently necessary component of the water quality based approach to limit and control the discharge of toxic substances. However, a direct ecological basis for such flow regulation is lacking and, furthermore, may not be relevant so that one flow duration determines ecological health and well-being. It is simplistic and ecologically unrealistic to expect that worst case biological community performance can only be measured under a $Q_{7,10}$ flow or some facsimile thereof.

There have been efforts to define ecologically critical flow thresholds using a toxicological rationale (USEPA 1986). This involved making judgements about the number, of exceedences of acute and chronic chemical criteria that could occur *without causing harm to the aquatic community.* This effort attempts to establish a minimum flow at which chemical and/or toxic unit limits could be set and not have the aquatic communities in "a perpetual state of recovery" (Stephan et al. 1985). While there has been no direct experimental validation of the maximum exceedence frequency using complex ecological measures in the ambient environment, validation efforts were later directed at using experimental streams (USEPA 1991f). These efforts, while being experimentally valid, retain many basic limitations inherent to surrogate criteria, one of which remains that a single species serves as "surrogates" for community health.

Establishing a single critical flow (i.e., $Q_{7,10}$, $Q_{30,10}$, etc.) on an ecological basis, however, is not only improbable under current science, it is technically inappropriate. There are simply too many additional variables that simultaneously affect the response and resultant conditions of aquatic communities both spatially and temporally. Some can be estimated (e.g., duration of exposure, chemical fate dynamics, and additivity), but many cannot because of the intensive data collection and analysis requirements; other phenomena are simply not adequately understood, yet their influence is integrated in the biological result.

The ecological ramifications of low-flow conditions (particularly extreme drought) in small streams has probably contributed much of the attention given to critical low'-flow. The results of low stream flow alone can be devastating in small watersheds (particularly those that have been modified via wetland destruction) during extended periods of severe drought (Larimore et al. 1957). The principal stressor in these cases is a loss of habitat via desiccation in which organisms either leave or die during these periods. Ironically enough, the sustaining flow provided by a point source discharge can mitigate the effects of desiccation if chemical conditions are minimally satisfactory for organism function and survival. While this may seem enigmatic in light of current strategies to regionalize wastewater flows, the presence of water with a seemingly marginal chemical quality can successfully mitigate what otherwise would be a total community loss. As was previously mentioned, this is dependent on the frequency, duration, and magnitude of any chemical stresses and local faunal tolerances. Small headwater streams (typically less than 10 to 20 mi^2 drainage areas) commonly experience near zero flows during extended dry weather periods, sometimes during several consecutive summers. Given the historical loss of wetlands that functioned to sustain flows during dry weather periods, strategies such is opting for small wastewater treatment plants instead of regionalization need to be considered if the aquatic communities in headwater streams are to be restored and maintained. In this situation, the discharge flow assumes the functional loss of the sustained dry weather flows formerly produced by wetlands. While this may seem contradictory the far worse consequences of repeated desiccation are far worse from a biological integrity standpoint.

Chemical-numerical applications necessarily have their basis in dilution scenarios. However, these types of simplified analyses are no match for the insights provided into the chemical, physical, and biological dynamics that are "included" in the condition of the resident biota. The resolution of steadystate chemical application techniques suffers when applied to extreme low-flow or high-flow conditions. Site-specific factors that outweigh the importance of flow alone include the availability and quality of permanent pools and other refugia, gradient, organism acclimatization, and riparian characteristics such as canopy cover. Together these and other factors determine the ability of a biological community to function and resist stress under worst case low-flow conditions and hence retain the essential elements of biological integrity. It would be a serious mistake to draw the conclusion that the only important function of stream flow is to dilute pollutant concentrations when in fact the influence on physical habitat, both flow and water volume, is a far more important factor. The misconceptions about the role of stream flow has not only hampered efforts to more accurately manage wastewater flows in small stream's, but in some cases has actually led to policies that have resulted in far more devastating ecological impacts than that experienced under the original problems. The most frequently cited concept is that biological data collected during any time other than $Q_{7,10}$ critical flow does not represent the effect of worst case conditions and therefore has limited applications in water quality based issues. Sampling under worst case, low-flow conditions is simply not necessary when measuring the condition of communities that have relatively long life spans and carry out all of their life functions in the waterbody. It is inappropriate to expect biological community condition (which is the integrated result of physical, chemical, and biological factors) to be so dependent on a temporal extreme of a single physical variable.

4.0 CRITICISMS OF THE INDEX OF BIOTIC INTEGRITY

Suter (1993) critically evaluated ecological health and IBI. Although he states that his "paper does not attack the concept [IBI] but rather the much more limited belief that the best way to use ... biosurvey data is to create an index of heterogenous variables [multimetric approach] and claim that it represents ecosystem health." The following is a list of his criticisms and a response to a potentially limited viewpoint of the IBI.

4.1 Ambiguity

Suter suggests that while using multimetric indices one cannot determine why values are high or low. The IBI utilizes multiple metrics to evaluate the water resource. One of the greatest advantages of IBI is that the site score can be dissected to reveal patterns exhibited at the specific reach compared to the reference community. Overall site quality can be determined from both the composite score and evaluation of each of the individual metrics. This reduces ambiguity compared to single metric indices such as the Shannon-Wiener Diversity Index.

4.2 Eclipsing

The eclipsing of low values of one metric can be dampened by the high values of another metric. Suter suggested that the density and disease linkage in epidemiology is and interrelated effect. He suggests that when toxic chemicals are involved, the disease factor may not be reflective of the density or quality of an otherwise unimpacted community. Studies by Ohio EPA (1987a,b) and other authors (e.g., Karr et al. 1985a,b; Karr et al. 1986) have shown that when the IBI is assessed properly each metric provides relevant information, which determines the position along a continuum of water resource quality. Thus, some sites may score well in some areas but poorly in other metrics depending on levels of degradation. Thus the reference condition is critical in determining the least impacted condition for the region.

4.3 Arbitrary Variance

Variance demonstrated in indices may be high due to the compounding of individual metric variances. Suter further suggests that other statistical properties of multimetric variables may be difficult to define. In studies conducted by Ohio EPA (Rankin and Yoder 1990, Yoder 1991b) they showed that IBI variability increased at highly degraded and disturbed sites but was low and stable at high-quality sites with increased biological integrity. The amount of variability within any of the component IBI metrics is irrelevant and does not necessarily have to be on the same scale assuming that proper metrics are selected and knowledge of how the metrics are applied are assessed by the field biologist. The high degree of resultant variability at sites that exhibit low biological integrity is an important indicator of site structure and function.

4.4 Unreality

Suter argues that using multimetric approaches results in values with “nonsense units.” He suggests that the IBI does not use “real” properties to describe the status of the reach specific water resource. In contrast, he used an example of dose response curves or habitat suitability to better predict a real-world property such as the presence of trout in a stream following a defined perturbation. In Suter’s simplistic approach to this complex problem he fails to recognize the multiple stresses that could potentially limit the possibility of aquatic organism uses of a stream. The water resource manager is not only interested in whether a species is present or absent from a stream reach, but also puts more weight on the species interactions in a web of dynamic interactions. The resultant hyperniche, defined by not only a single species but multiple species, becomes impossible with the limited amount of chemical specific information available. Likewise, the modeling of the synergistic and additive effects of multiple stressors suggests that the IBI is only sensitive to toxic influences. This has shown to be a poor assumption since the IBI can determine poor performance from point source, nonpoint source, and combinations of these effects. The IBI does use real-world measures, which individually are important attributes of a properly functioning and stable aquatic community. Additionally, the assessment of the aquatic community is enhanced by the acquisition of appropriate habitat information. It is highly recommended that all assessments include not only biological community information but habitat information (Davis and Simon 1988).

4.5 Post hoc Justification

Suter suggests that the reduction in IBI values is a tautology since the assessment of poor biological integrity is a result of the reduced score. He further suggests that the IBI will only work if all ecosystems in all cases become unhealthy in the same manner. The IBI metrics are a priori assumed to measure a specific attribute of the community. Each metric is not an answer unto itself and not all measure only attributes of a properly functioning community (e.g., percent disease). The metric must be sensitive to the environmental condition being monitored. The definition of degradation responses a priori is justified if clear patterns emerge from specific metrics. Although the probability of all ecosystems becoming unhealthy in the same manner is unrealistic it is important to note that response signatures are definable (Yoder and Rankin, Chapter 17) based on patterns of specific perturbations.

4.6 Unitary Response Scales

Suter suggests that combining multimetric measures into a single index value suggests only a single linear scale of response and, therefore only one type of response by ecosystems to disturbance. Suter fails to recognize that the individual patterns exhibited by the various individual metrics usually reduces to single patterns in the community. For, example, whether discussing siltation, reduced dissolved oxygen, or toxic chemical influences all reduce the sensitive species component of the community and reduces species diversity. Thus, although multiple measures of the individual metrics results in multiple vectors explaining those dynamic patterns, a priori predictions of the metric response will result in the biological integrity categories defined by Karr et al. (1986).

4.7 No Diagnostic Results

Suter suggests one of the most important uses of biological survey data is to determine the cause of changes in ecosystem properties. He further suggests that by combining the individual metrics into a single value causes a loss of resolution when attempting to diagnose the responsible entity. This is the same argument raised in the ambiguity discussion above. The greatest use of IBI is the ability to discern differences in individual metrics and determine cause and effect using additional information such as habitat, chemical water quality, and toxicity information. The inverse, however, is not apparent when attempting to reduce chemical water quality and toxicity test information into simple predictions of biological integrity based on complex interactions.

4.8 Disconnected from Testing and Modeling

Suter suggests that the field results determined from the IBI need to be verified in the laboratory using controlled studies such as toxicity tests. This is a narrow viewpoint of the complex nature of the multimetric approach. Seldom does the degradation observed at a site result from a single chemical contaminant. To suggest that a single-species or even multiple-species (usually run individually) toxicity test can predict an IBI is ridiculous given that the effects of siltation, habitat modification, guild and trophic responses cannot be adequately determined in a laboratory beaker. Those aspects of a community that can be tested in the laboratory has validated the individual metric approach, i.e., thermal responses. It is the compilation of the various attributes that gives IBI a robust measure of the community.

4.9 Nonsense Results

Suter indicates any index based on multiple metrics can produce nonsense results if the index has no interpretable real-world meaning. Suter suggests that green sunfish (considered a tolerant species in the IBI) may have a greater sensitivity to some chemicals than some “sensitive species,” and that the reduction of these contaminants may enable increases in green sunfish populations which result in a reduction in biological integrity. However, Suter has mistakenly suggested that green sunfish have a greater position in the community than do sensitive species. Green sunfish and other tolerant species are defined by the species ability to increase under degraded conditions (Karr et al. 1986; Ohio EPA 1987a,b). Range extensions and the disruption of evenness in the community often occurs at the expense of other sensitive species. This suggests that scoring modifications and other mechanisms for factoring out problems when few individuals are collected are not real-world situations.

4.10 Improper Analogy to Other Indices

Since environmental health as a concept has been compared to an economic index several authors have argued that the environmental indices are not generally comprehensible and require an act of faith to make informed judgements or decisions. The IBI has greatly improved the decision-making process by removing the subjective nature of past biological assessments. By using quantitative criteria (biological criteria) to determine goals of the Clean Water Act (attainable goals and designated uses) the generally comprehensible goals of the IBI enable a linkage between water resource status and biological integrity. This does not require an act of faith; rather, it broadens the tools available to water resource managers for screening waterbody status and trends.

5.0 DEVELOPMENT OF METRIC EXPECTATIONS

The IBI requires quantitative expectations of what a fish community should look like under reference or least impacted conditions (Karr et al. 1986; USEPA 1990a). Each metric has its own set of expectations, and metric expectations often vary with ecosystem size or location (e.g., Fausch et al. 1984). Generating an acceptable set of expectations is perhaps the most difficult part of developing a new version of the IBI or effectively applying an existing version to a new geographic area. Usually, expectations have been developed on a watershed or regional basis, and have been derived from recent field data (Hughes et al. 1986, 1990; Plafkin et al. 1989).

Because fish communities may differ substantially between different geographic areas, accurate delineation of appropriate regions for development and application of expectations is critical. Early efforts defined regions based on watershed boundaries (Karr 1981; Fausch et al. 1984; Moyle et al. 1986; Karr et al. 1986), recognizing the major faunal differences that may exist among drainage basins (Hocutt and Wiley 1986). More recently, many IBI versions have used Omernik’s (1987) ecoregions as their geographic framework for setting expectations (Hughes and Gammon 1987; Geise and Keith 1989; Langdon 1989; Ohio EPA 1987a,b; Simon 1991; Larsen and Christie 1993). Fish community composition in streams has been shown to differ among ecoregions (Larsen et al. 1986; Hughes et al. 1987; Rohm et al. 1987; Whittier et al. 1988; Lyons 1989; Hawkes et al. 1986), and several workers have argued that ecoregions may be more appropriate than drainage basins in developing regional expectations (Hughes et al. 1986; Hughes and Larsen 1988; Gallant et al. 1989; Omernik and Griffith 1991; Omernik, Chapter 5). Some IBI versions have used a combination of ecoregion and watershed boundaries to delineate regions (Fisher 1989; Simon 1991; Lyons 1992). Regardless of which regional framework is used the final tuning should utilize statistical multivariate approaches to determine patterns that may not follow any prescribed regional framework such as that observed among forested regions throughout the midwest (Fausch et al. 1984) and for large and great rivers in Indiana (Simon 1992).

Two general approaches have been used to generate quantitative metric expectations for a particular geographic area. The first approach requires identification and sampling of a limited number of representative sites in relatively undegraded or least impacted ecosystems (Hughes et al. 1986; Gallant et al. 1989; Warry and Hanau 1993). Hughes (1985) and Hughes et al. (1986) provided detailed guidelines for selecting appropriate least impacted sites. Data from these sites are then used to define expectations and establish metric scoring criteria. This approach has been used successfully with stream fish communities in Ohio (Larsen et al. 1986, 1988; Whittier et al. 1987; Ohio EPA 1987a,b), Arkansas (Rohm et al. 1987; Geise and Keith 1989), Vermont (Langdon 1989), and New York (Bode and Novak 1989).

The second approach does not involve delineation of specific high-quality or least impacted sites, but require more data. Under this approach, a large number or sites are surveyed in a systematic fashion to provide a representative view of the region. The best values observed for each metric, even if they do not come from the highest quality sites, are then used to define the expectations and set scoring criteria. This approach, which has been widely used (e.g., Karr 1981; Fausch et al. 1984; Karr et al. 1986; Moyle et al. 1986; Hite and Bertrand 1989; Maret 1989; Simon 1991; Lyons 1992; Osborne et al. 1992), has worked best either when it has been difficult or impractical to identify least impacted sites, or when a large database has already been present, but additional data collection has not been possible.

Within a particular geographic area, different metrics, expectations, and scoring criteria have been used for different types of ecosystems. Existing versions of the IBI have typically recognized several different types of lotic ecosystems, although the distinction between them has not always been clear and has varied among versions. Ohio EPA (1987a,b) distinguished among headwater streams, wadable streams and rivers, and boat-sampled rivers based on the expected fish fauna, the size of the drainage basin, and the fish sampling techniques that was most appropriate. Simon (1992) separated large and "great" rivers based on drainage basin area. Lyons (1992) discussed the differences in fish fauna, response to environmental degradation, and maximum summer water temperatures between Wisconsin warmwater and coldwater streams and concluded that a single version of the IBI would not be possible. For lentic ecosystems, the primary distinction has been between impoundments and natural lakes, although within each ecosystem type it seems that size- and temperature-based stratification are warranted (Hughes et al. 1992; Larsen and Christie 1993). Work on Tennessee Valley Authority impoundments has indicated that reservoirs with different hydrologic regimes and watershed position (mainstem, constant water level vs. tributary, fluctuating water levels) need to be treated separately and that different relative locations within large impoundments require different metric expectations (Dionne and Karr 1992).

Expectations for species richness metrics have often been an increasing function of ecosystem size (Fausch et al. 1984, 1990; Karr et al. 1986; Miller et al. 1988). For streams and rivers, ecosystem size has been expressed as stream order (Karr et al. 1986), drainage basin area (Ohio EPA 1987a,b), or mean channel width (Lyons 1992). Each of these measures has its pros and cons (Hughes and Omernik 1981, 1983; Lyons 1992). For impoundments and lakes, surface area has been used (Larsen and Christie 1993). Typically, the size vs. species richness function is assumed to be either log-linear or asymptotic or a combination of both, and is fit graphically by eye such that about 5% of the data points fall above the resulting maximum species richness (MSR) line. The by-eye method is imprecise, and often numerous but equally valid MSR lines could fit to the same data set. Some workers constrain the intercept to be at the graphical origin; however, since fish are found in even the smallest size streams problematic slopes result in the smallest headwater streams or the largest river. The MSR line should not be constrained to go through the origin (J. R. Karr, personal communication). Lyons (1992) developed an objective graphical technique for drawing MSR lines, but generally a more precise statistical procedure for generating MSR lines would be desirable. It is important to recognize what information is supported by the data and not to over extend the results beyond what can be supported.

Several other factors have been considered in the development of metric expectations. In some streams and rivers, gradient strongly influences fish species richness and composition, with high-gradient sites often having lower richness and different species than nearby low-gradient sites (Leonard and Orth 1986; Miller et al. 1988; Lyons 1989). Recent studies indicate that the relative position of ecosystems within the drainage network can also affect species richness. Fausch et al. (1984), Karr et al. (1986), and Osborne et al. (1992) demonstrated that small adventitious streams (Gorman 1986) that flowed directly into much larger rivers had higher species richness values and IBI scores than environmentally similar headwater streams that were distant from large rivers. They concluded that small adventitious and headwater streams needed different metric expectations. Lyons (1992) found that the number of sunfish species was higher in streams near lakes and large rivers than in those distant, and developed separate MSR lines for each case. Although drainage position has received only general, larger-scale consideration thus far in IBI versions for lentic habitats (Dionne and Karr 1992), the presence of tributaries and the relative "connectedness" of lakes with other bodies of water is known to have a major influence on lake fish community composition (summarized by Tonn et al. 1990).

6.0 MAJOR IBI MODIFICATIONS

6.1 Ichthyoplankton Index

As a rule, existing versions of the IBI have focused on juvenile and adult fish, and explicitly excluded larval and small young-of-the-year fish. For example, Karr et al. (1986) recommended against inclusion of fish under 20 mm total length, whereas Lyons (1992) excluded all fish below 25 mm. Angermeier and Karr (1986) argued that inclusion of any size of young-of-the-year fish would reduce the accuracy of the IBI applications. The difficulties in sampling and identifying larvae coupled with the commonly high temporal variability in larval abundance have been the main arguments against the use of larvae and young-of-the-year fish in the IBI.

However, Simon (1989) developed a version of the IBI, termed the Ichthyoplankton Index, specifically for larval fish in lotic habitats. The early life history stage of fishes have been recognized as the most sensitive and vulnerable life stage (Blaxter 1974; Moser et al. 1984; Wallus et al, 1990), and, thus, would be particularly sensitive to certain types of ecosystem degradation.
The Ichthyoplankton Index is organized into 11 metrics covering taxonomic composition, reproductive guild, abundance, generation time, and deformity categories (Table 3). Since much of the North American fauna is incompletely described (Simon 1986), a family-level approach was designed to evaluate early life history stages. Reproductive guilds are based on Balon (1975, 1981). Simon (1989) assigned tolerance values based on sensitivity to siltation, sediment degradation, toxic substances, and flow modifications. Larvae of Clupeidae, Sciaenidae, and Osmeridae were excluded from metric calculations because they often reached such high abundances that their inclusion would obscure abundance patterns for other taxa.

The Ichthyoplankton Index has not been field tested, so little discussion of precision and accuracy is possible. Simon (1989) felt that its greatest value would be in assessing the integrity of nursery habitats, particularly backwaters of large rivers. The primary limits to its widespread and effective use will likely he the difficulty of collecting the necessary data to establish specific regional expectations and scoring criteria, and of establishing a standardized sampling period and technique.

Table 3. List of Original Ichthyoplankton Index Metrics Proposed by Simon (1989) for Streams in the United States Based on Early Life History and Reproductive Biology Characteristics

Taxonomic composition metrics
1. Total Number of Families
2. Number of Sensitive Families
3. Equitability/Dominance
4. Family Biotic Index

Reproductive guild metrics
5. Percent Nonguarding Guild
6. Percent Guarding Guild
7. Percent Bearing Guild
8. Percent Simple Lithophil Guild

Abundance, generation time, and deformity metrics
9. Catch per Unit Effort
10. Mean Generation Time
11. Percent Deformity or Teratogenicity

6.2 Warmwater Streams

The majority of IBI applications have involved small- to medium-sized wadable warmwater streams and numerous modifications have been made to the original version to reflect regional or ecosystem differences in fish communities. Some of the original metrics, have been changed in nearly every subsequent version, whereas others have been largely retained (Table 2).

6.2.1 Species Richness and Composition

Changes in the original species richness and composition metrics have typically been limited, most commonly involving expansion of taxonomic groups to include other species. This has been done when the original taxonomic group had few species in the area or ecosystem type of interest. For example, the number of catostomid species metric has been expanded to include Ictaluridae in Ontario (Steedman 1988), all benthic insectivorous species in the northeastern United States (Miller et al. 1988), or all benthic species in France (Oberdorff and Hughes 1992). The number of darter species metric has been broadened to encompass Cottidae in Ohio headwaters (Karr et al. 1986; Ohio EPA 1987b) and in Ontario (Steedman 1988), or Cottidae and *Noturus* species (Ictaluridae) in northwestern Indiana and northern Wisconsin (Simon 1991; Lyons 1992), and has been replaced with the number of Cyprinidae species in Arkansas (Karr et al. 1986), Oregon (Hughes and Gammon 1987), Colorado (Schrader 1986; Bramblett and Fausch 199 1), Ohio headwaters (Ohio EPA 1987a,b), Indiana headwaters (Simon 1991), and Minnesota headwaters (Bailey et al. 1993), or proportion of cyprinids with subterminal mouths in the Red River of the North basin (Goldstein et al. 1994). The number of sunfish species metric has been expanded to include all water column (nonbenthic) species in the northeastern United States (Miller et al. 1988) and in France (Oberdorff and Hughes 1992), and substituted by the number of headwater species (species generally restricted to good quality permanent headwaters) in Ohio and Indiana headwaters (Ohio EPA 1987a,b; Simon 1991) and the Red River of the North basin (Goldstein et al. 1994).

6.2.2 Indicator Species Metrics

Metrics proposed by Karr (1981) and Karr et al. (1986) to evaluate species sensitivity to human influence on watersheds have been the most frequently changed IBI metrics (Table 2). Usually the changes have resulted from differences of opinion in species sensitivity, differences in drainage area relationships, and use of other more regionally representative species to act as tolerant species surrogates. Oberdorff and Hughes (1992) substituted the number of intolerant species with the presence of northern pike.

The most frequently changed metric among IBI efforts has been the proportion of green sunfish. Green sunfish are usually abundant only in small creeks and streams and do not reflect deteriorating habitat in wadable rivers and larger rivers. Green sunfish are also restricted largely to the central United States, which makes them a poor choice for widespread application in other geographic areas. The proportion of green sunfish has been replaced with more regionally appropriate tolerant species, e.g., common carp (*Cyprinus carpio*), longear sunfish *(Lepomis megalotis*), creek chub (*Semotilus atromaculatus*), white sucker (*Catostomus commersoni*), blacknose dace *(Rhinichthys atratulus*), or roach (*Rutilus rutilus*) (Leonard and Orth 1986; Schrader 1986; Hughes and Gammon 1987; Miller et al. 1988; Oberdorff and Hughes 1992), or many versions have included all tolerant species to increase the sensitivity of this metric. Some have replaced the proportion of green sunfish with either the proportion or number of exotic species (Schrader 1986; Hughes and Gammon 1987; Crumby et al. 1990; Fisher 1990; Bramblet and Fausch 1991). Headwater streams typically do not provide sufficient habitat for many of the tolerant species, so a substitution of the metric includes the proportion of pioneer species in Ohio and Indiana (Ohio EPA 1987a,b; Simon 1991). Pioneer species are small tolerant species that are the first to recolonize headwaters following desiccation or fish kills (Smith 1971). Goldstein et al. (1994) suggests that tolerance is too subjective, so this metric may be substituted with evenness.

6.2.3 Trophic Function Metrics

The trophic composition metrics have been relatively unchanged across widespread geographic application of the IBI (Table 2). The omnivore and carnivore metrics have required the fewest modifications. Leonard and Orth (1986) substituted the proportion of generalized feeders because they found that few species in their streams fit the omnivore definition. Filter-feeding species were not considered omnivores for this metric nor were species that had a generally plastic response to diet when confronted with degraded habitat conditions, such as creek chub and blacknose dace. The proportion of individuals as insectivorous cyprinids has been modified in most versions to include all specialized invertebrate feeders or all insectivores. These broader groupings have proven generally more sensitive, particularly where the species richness of insectivorous cyprinids is low (Miller et al. 1988). The carnivore metric has been deleted from some analyses due to lack of occurrence (Leonard and Orth 1986; Fausch and Schrader 1987) or due to drainage area relationships (Ohio EPA 1987a,b; Simon 1991). Hughes and Gammon (1987) suggested that the primary carnivore in the Willamette River, Oregon, was an exotic and tolerant of degraded conditions indicating low integrity when abundant. Goldstein et al. (1994) substituted the percent of various trophic categories with the percent biomass of trophic guilds.

6.2.4 Reproductive Function Metrics

The proportion of hybrids metric has been difficult to apply in most geographic regions. The hybrid metric was designed to reflect tendencies for breakdown in reproductive isolation with increasing habitat degradation. Other researchers have not found hybridization to be correlated with habitat degradation (Pflieger 1975; Ohio EPA 1987a,b). Problems with field identification, lack of hybrids even in some very degraded habitats, and presence of hybrids in some high quality streams among some taxa have precluded this metric from being successfully applied. In western fish faunas, this metric has been modified with the number of introduced species (Fausch and Schrader 1987; Hughes and Gammon 1987). Courtney and Hensley (1980) consider the increase of normative species as a form of biological pollution which is similar to the original intent of the metric.

In some geographic areas the hybrid metric has been deleted from the index and replaced with the proportion of simple lithophils (Table 2), which are defined as those species spawning over gravel without preparing a nest or providing parental care. The proportion of simple lithophilous spawning species is believed to be inversely correlated with habitat degradation based on the destruction of high-quality spawning habitat (Berkman and Rabeni 1987; Ohio EPA 1987a,b; Simon 1991). Berkman and Rabeni (1987) found an inverse relationship between the number of lithophilous spawning species and siltation. Goldstein et al. (1994) utilized the ratio between broadcast spawning and nest-building cyprinids to determine reductions in substrate quality in the species depauperate Hudson River drainage.

6.2.5 Abundance Metrics

Catch per unit of effort (CPUE) has been retained in most versions of the IBI; however, some have excluded tolerant species (Ohio EPA 1987b; Lyons 1992). CPUE is influenced by stream size and sampling efficiency. Karr et al. (1986) suggested using relative CPUE to set scoring criteria for the total number of individuals metric; however, fish density and sampling efficiency tends to decrease as watershed and stream size increases (Thompson and Hunt 1930; Larimore and Smith 1963; Miller et al. 1988). Steedman (1988), Gammon (1990) and Lyons (1992) found abundance to be higher at moderate levels of degradation (i.e., nutrient enrichment) and lowest at severe levels. Lyons (1992) used the abundance metric as a correction factor for the overall IBI score only when very low CPUE was observed. Hughes and Gammon (1987) and Fisher (1990) included biomass and estimates for fish and amphibians, and the density of macroinvertebrates to increase sensitivity in their streams that had relatively low species diversity.

6.2.6 Fish Condition Metrics

The percent of diseased and deformed individuals has been retained in most versions of the IBI. However, infestation by parasites and protozoans has often been eliminated from this metric due to a lack of correlation between parasite burden and environmental quality (Whittier et al. 1987; Steedman 1991). Most researchers have included only obvious external anomalies such as deformations, eroded fins, lesions, or tumors. Lyons (1992) used this metric as a correction factor for the overall IBI score only when a relatively high percentage of diseased and deformed individuals were present at the sample site.

6.3 Large Rivers

Relatively little work has been directed towards modifying the IBI for use on rivers too large to sample by wading. Difficulties in accurately and easily sampling fish assemblages and the scarcity of appropriate least impacted reference sites for setting metric expectations have hampered IBI development for these types of systems. Only five published versions are currently available, covering the Willamette River in northwestern Oregon (Hughes and Gammon 1987), the large rivers of Ohio (Ohio EPA 1987a,b), the Seine River in north-central France (Oberdorff and Hughes 1992), the Current and Jacks Fork Rivers in southeastern Missouri (Hoefs. and Boyle 1992), and the large rivers of Indiana (Simon 1992). The Indiana version distinguishes between "large" rivers, with drainage areas of 1000 to 2000 mi^2 , and "great" rivers, with drainage areas greater than 2000 mi^2 (Figure 1).

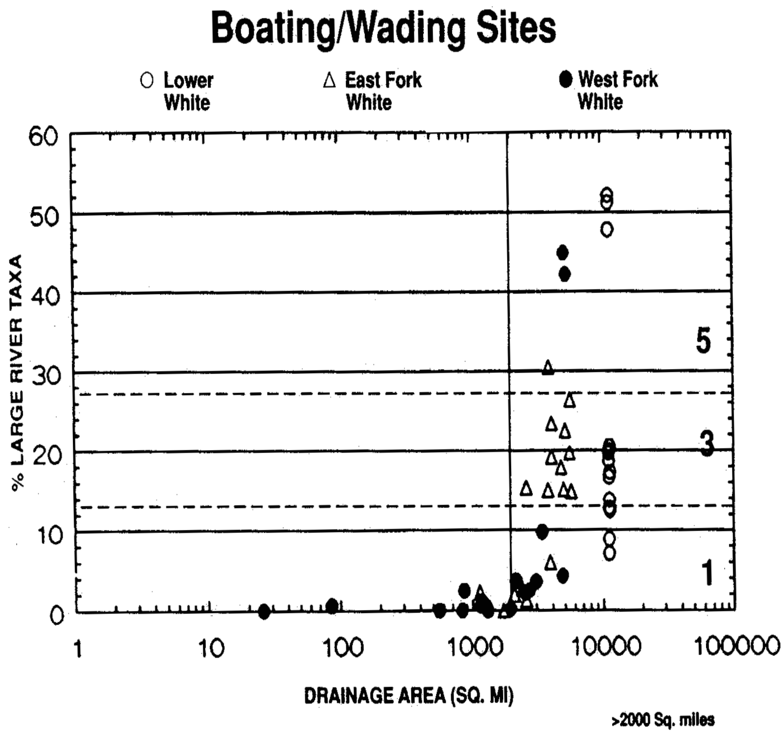

Figure 1. Maximum Species Richness (MSR) line plot of the proportion of large river taxa with drainage area. Drainage areas greater than 2000 mi^2 define the difference between large and great river metrics. (From Simon, T. P. 1992. *Development of Biological Criteria for Large Rivers with and Emphasis on an Assessment of the White River Drainage, Indiana.* USEPA, Chicago, Illinois.)

Generally, all four large river versions have metrics that are similar to those in IBI versions developed for smaller, wadable streams and rivers (Table 2). However, the Ohio, Missouri, and Indiana versions have incorporated two unique metrics especially tailored for larger rivers, the percent of individuals that are round-bodied catostomid species *(Cycleptus, Erimyzon, Hypentelium, Minytrema, and Moxostoma),* and the number of specialized large river species, which tend to be restricted to high-quality large rivers. Round-bodied catostomids are most common in relatively undegraded large rivers of eastern North America, and are sensitive to water pollution, thermal loadings, and habitat degradation (Gammon 1976; Karr et al. 1986; Ohio EPA 1987a,b; Simon 1992). Although round bodied in shape, the white sucker is not included in the round-bodied group because it is highly tolerant of poor water and habitat quality. Other common large-river catostomids, the carpsuckers *(Carpiodes)* and the buffaloes *(Ictiobus),* which are more laterally compressed, also are not included because they are omnivorous species that can survive in thermally stressed and degraded habitats.

The number of specialized large-river species metric has thus far been used only for great rivers in Indiana (Figure 1), although it seems a promising metric for large rivers in many different regions. This metric is based on studies in Missouri by Pflieger (197 1), where a fish fauna characteristic of high-quality reaches of large rivers was identified. A similar fauna has also been documented for Indiana (Gerking 1945), Illinois (Smith et al. 1971), Arkansas (Matthews and Robison 1988), and Kentucky (Burr and Warren 1986).

6.4 Modification of the IBI for Coldwater and Coolwater Habitats

Coolwater streams have a mean maximum daily temperature between 22 and 24°C during a normal summer and coldwater streams normally have maximum daily means below 22°C, whereas warmwater streams exceed 24°C (Lyons 1992). High-quality coldwater streams are dominated by salmonid and cottid species. High-quality coolwater streams are often too warm to support large populations of salmonid and cottids, but too cold to support the full complement of species found in warmwater streams. Cool and coldwater streams are common in much of Canada and the northern and western United States.

Most versions of the IBI have been developed for warmwater streams, rather than cool- or coldwater streams. Most applications of the IBI for coolwater and coldwater streams have been for high-gradient areas of the eastern and western United States (Leonard and Orth 1986; Moyle et al. 1986; Hughes and Gammon 1987; Langdon 1989; Steedman 1988; Fisher 1990; Oberdorff and Hughes 1992), and versions developed by Leonard and Orth (1986) and Miller et al. (1988) for the eastern United States were for predominantly coolwater systems. Steedman (1988) and Lyons (1992) demonstrated that warmwater versions of the IBI were inappropriate for use in coldwater streams in the north-central United States and Canada. In this region, degraded cool- and coldwater streams often show increased species richness for many groups of fishes, the opposite of what normally occurs in warmwater streams.

The general trend for coolwater and coldwater versions of the IBI has been for a reduction in the number of metrics, e.g., 8 to 10 for coldwater (with the exception of Hughes and Gammon 1987, which has 12 metrics but is both a warm and coldwater version), vs. 10 to 12 metrics for warmwater. This reduction in metrics reflects the simplified structure and function of coldwater fish communities relative to warmwater communities. Moyle et al. (1986) and Fisher (1990) also included the number of amphibian species and the density of macroinvertebrates in their coldwater versions.

6.5 Estuaries

Very little work has been completed in the development of estuarine versions of the IBI. Only one ocean estuary version is currently available. Thompson and Fitzhugh (1986) made significant changes in the IBI to reflect estuarine/marine component of the Louisiana fauna that they studied, but they retained the general framework of the index. Apparently, little additional testing and application of this version has taken place and many questions remain about sampling, natural dominance of certain species, and high inherent temporal/ seasonal variation in assemblage structure.

Preliminary investigations by Thoma (1990) on Lake Erie estuaries and Simon et al. (1989) for the Grand Calumet River and Indiana Harbor Canal are the only freshwater efforts. Thoma (1990) used Ohio EPA boat versions of the IBI and suggested replacement of the number of sunfish, number of sucker, proportion of round-bodied suckers, and proportion of simple lithophils with the proportion of exotics, number of vegetation associated species, and number of Lake Erie species. Simon et al. (1989) eliminated transient migratory species from their analysis in scoring the total number of species and proportional metrics.

6.6 Lakes

Little work has been published for lakes and reservoirs, but significant efforts are ongoing as part of the USEPA effort to develop consistent bioassessment protocols (Gerritsen et al. 1984). The USEPA's Environmental Monitoring Assessment Program (EMAP) is testing potential metrics on natural lakes in the northeastern United States (Hughes et al. 1992; Larsen and Christie 1993), and Karr and the Tennessee Valley Authority (TVA) (Dionne and Karr 1992; Dycus and Meinert 1993; Jennings, Karr and Fore, personal communication) are conducting research on IBI application in impoundments. Proposed metrics for lentic systems are listed in Table 4.

At present, EMAP has proposed eight potential metrics for evaluating the biotic integrity of natural lakes (Table 4; Hughes et al. 1993; Larsen and Christie 1993). Six of these metrics are the same as those used in some stream versions of the IBI, but two are unique, the overall age/size structure and the percentage of individuals above a certain size for selected species. These two metrics are designed to assess the reproductive success, survival, and vulnerability to increasing degradation for populations of key indicator species. Additional work is needed to more precisely define and set expectations for these two metrics, and to test their usefulness in field applications.

6.7 Impoundments

The original TVA version of the IBI for use in impoundments had 14 metrics (Table 4; Dionne and Karr 1992). Most of these metrics were taken from stream and river versions of the IBI, reflecting the intermediate nature of reservoirs between rivers and lakes. Six metrics were new: the number of small cyprinid and darter species, the percent of individuals as young-of-the-year shad *(Dorosoma cepedianum* and *D. petenense)* and bluegill *(Lepomis macrochirus),* the percent of individuals as adult shad and bluegill, the percent of species as plant and rock substrate spawners, the number of migratory spawning species, and the fish health score for largemouth bass *(Micropterus salmoides).* The number of small cyprinids and darters metric was based on the high species richness for these taxa that included to assess the quality of littoral zone and tributary spawning areas, which can be degraded by excessive sedimentation, fluctuations in water level, and direct physical modification. The physical health score represented an effort to objectively evaluate the physical condition and physiological state of an important reservoir species through a thorough external and internal examination. The procedure followed was based on a scoring system developed by Goede (1988).

More recently, a second TVA reservoir version of the IBI has been developed, termed the Reservoir Fish Assemblage Index (RFAI; Jennings, Karr, and Fore, personal communication). The RFAI has a somewhat different set of 12 metrics (Table 4), with the changes in metrics designed to improve sensitivity to environmental degradation and to increase adaptability to different types of reservoirs. However, results from applications of both the original TVA version and the newer RFAI have often not accurately reflected what are believed to be the true patterns in environmental health within and among reservoirs, and additional modifications will probably be necessary to develop better versions of the IBI for impoundments (Jennings, personal communication). In the RFAI, the number of sunfish species has been substituted for the number of small cyprinid and darter species metric, the percent of individuals as invertivores has been substituted for the percent of individuals as specialized benthic insectivores metric, and the number of simple lithophilous spawning species has been substituted for the percent of individuals as plant and rock substrate spawners metrics. Three metrics have been dropped: the percent of individuals as piscivores, the percent of individuals as young-of-year, and percent adult shad and bluegill. Shad also are not included in the calculation of the four remaining percent abundance and catch-per-effort metrics. The physical health score metric has thus far been retained in the RFAI, but it has been recommended for deletion because of its low sensitivity to known environmental problems.

Table 4. List of Index of Biotic Integrity Metrics Proposed for TVA Reservoirs by Dionne and Karr (1992) (R1) and by Jennings, Karr, and Fore (personal communication) (R2) and for Natural Lakes in the Northeastern U.S. by Hughes et al. (1992) (L).

Species Richness And Composition Metrics

1. Total Number of Species (R1, R2, L)
2. Number of Catostomidae Species (Rl)
3. Number of Small Cyprinidae and Darter Species (R1)
4. Number of Sunfish Species (R2)

Indicator Species Metrics

5. Number of Intolerant Species (R1, R2, L)
6. Age and/or Size Structure for Populations of Selected Species (L)
7. Percent of Individuals Larger than a Certain Size for Selected Species (L)
8. Percent of Individuals that are Tolerant Species (R1, R2, L)
9. Percent of Individuals that are Exotic Species (L)

Trophic Function Metrics

10. Percent of Individuals that are Specialized Benthic Insectivores (R1)
11. Percent of Individuals that are Invertivores (R2)
12. Percent of Individuals that are Omnivores (R1, R2)
13. Percent of Individuals that are Piscivores (R1)
14. Percent of Individuals that are Young-of-Year Shad and Bluegill (R1)
15. Percent of Individuals that are Adult Shad and Bluegills (R1)

Reproductive Function Metrics

16. Percent of Individuals that are Plant and Rock Substrate Spawners (R1)
17. Percent of Individuals that are Simple Lithophilous Species (R2)
18. Percent of Individuals that are Migratory Spawning Species (R1, R2)

Abundance and Condition Metrics

19. Abundance or Catch per Effort of Fish (R1, R2, L)
20. Percent of Individuals that are Diseased, Deformed, or Have Eroded Fins, Lesions, or Tumors (R2, L)
21. Fish Health Score for Largemouth Bass (R1, R2)

Metric expectations for the two TVA versions of the IBI differ between mainstem and tributary reservoirs and between the inflow, transition, and forebay (near dam) regions of individual reservoirs (Dionne and Karr 1992; Jennings, Karr, and Fore, personal communication). These differences in expectations reflect inherent differences in fish assemblages within and among reservoirs. Development of metric expectations for the TVA versions has been hampered by the high spatial and temporal variability typical of reservoir fish assemblages, the limited range of environmental quality and high interconnectedness among the reservoirs studied, and the lack of least impacted reference systems for comparison (Jennings, Karr, and Fore, personal communication). Generally, it will be difficult to define and identify appropriate least impacted reference waters for reservoirs, as reservoirs are, by nature, highly impacted and artificially modified rivers.

7.0 FUTURE DEVELOPMENTS

New versions of the IBI continue to be developed for wadable warmwater streams. To our knowledge, efforts are currently underway to generate IBI versions for streams in the coastal plain of Maryland and Delaware, the New River drainage in West Virginia, the Ridge and Valley Physiographic Province of the central Appalachian Mountains, Mississippi River tributaries in northwestern

Mississippi, the Minnesota River drainage in central Minnesota, and the Red River of the North drainage in northwestern Minnesota and northeastern North Dakota. It is likely that we are unaware of additional ongoing efforts. Despite the large amount of current and past work on IBI versions for wadable warmwater streams, much remains to be done. In particular, versions are needed for most of Canada, the species-rich southeastern United States, and the species-poor western United States. Very small and intermittent streams have received insufficient attention everywhere but Ohio and Indiana. Low-gradient, wetland streams are another under represented habitat type. Finally, and perhaps most importantly, many existing versions have as yet not been property validated with independent data. The development of an IBI version is usually an iterative, somewhat circular process, and each new version requires a thorough field test and critique on a new set of waters before it can be safely applied.

For those areas with validated versions of the IBI, the next challenge will be to incorporate the IBI into routine monitoring and assessment programs. Many versions still remain research-level tools that are not widely used in the region where they were developed. Although research has shown them to be useful and reliable approaches to assess resource condition, many states have not moved forward to implement them. However, this will change in the United States as the USEPA encourages and mandates development and application of biocriteria into water quality standards programs. Currently, Ohio EPA is the best example of how the IBI can be incorporated into a state water resource management and protection program.

Large rivers are also the subject of increasing efforts, although considerable work is still required. In the upper Mississippi River, the U.S. Fish and Wildlife Service's Long Term Resource Monitoring Program is attempting to develop an IBI version using existing large standardized data sets from multiple information sources. However, undesirable statistical properties of this version have stymied further development and application (Gutreuter, Lubinski, and Callen, personal communication). Significant unresolved issues include the determination of appropriate sampling techniques, temporal and spatial scales, and scope of sampling effort. Clearly, multiple sampling techniques will be needed to get a representative snapshot of the large-river fish community, but there are problems with aggregating results from different gears/techniques and amounts of effort. These are long-identified problems, and they await resolution before large-river IBI development efforts can move forward. Ohio River Valley Sanitation Commission's (ORSANCO) Biological Water Quality Subcommittee and Ohio EPA are developing standard operating procedures on the Ohio River that will enable representative sample collection and data interpretation.

Authors Lyons and Simon are independently working with colleagues on IBI versions for low- to moderate-gradient coldwater systems in the north-central United States. At present, the Lyons version is a greatly simplified warmwater IBI, focusing on the presence and relative abundance of obligate coldwater species, intolerant species, and tolerant species. The Simon version has more metrics, but also utilizes aspects of warmwater versions. Based on findings of Lyons (1992), Simon proposes a "reverse" scoring system, where greater numbers of species result in a lower metric score, the opposite of scoring procedures for warmwater versions. IBI streams versions are also being developed for areas outside North America. Lyons and colleagues are working on versions for streams in west-central Mexico, and Ganasan and Hughes (personal communication) are working on a version for central India.

Few new developments are occurring with estuarine IBI's. It appears that Thompson (Thompson and Fitzhugh 1986) never developed his Louisiana version any further. However, Thoma (1990) is in the process of validating his version for Lake Erie. Further efforts are needed to enable use of IBI's over broad geographic areas in other Great Lake and coastal estuaries.

Lake and impoundment IBIs are being improved, but have really only just begun. Natural lake MI work is still concentrated on the preliminary stages of determining appropriate sampling methodologies, identifying least impacted sites, developing databases, and exploring the properties of potential metrics. The work is largely being done through EMAP, but up until now has focused exclusively on the northeastern United States.

Biocriteria for impoundments are being developed by TVA incorporating the RFAI into broader monitoring program. Dycus and Meinert (1993) have utilized dissolved oxygen, sediment, algae, benthic macroinvertebrates, and bacteria indices along with the RFAI to generate an overall picture of ecosystem health for reservoirs. However, Jennings, Karr, and Fore (personal communication) feel that the RFAI still requires significant work before it can be broadly applied. In particular, more data from additional interconnected reservoirs that range from poor to excellent in environmental quality are needed to improve the RFAI and ensure sensitivity to a wide range of environmental conditions.

8.0 SUMMARY

The Index of Biotic Integrity has been a widely applied and effective tool for using fish assemblage data to assess the environmental quality of aquatic habitats. The original version of the IBI has been modified in numerous ways for application in many different regions and habitat types, and the IBI is now best thought of as a family of related indices rather than a single index. The commonalities linking all IBI versions are a multimetric approach that rates different aspects of fish community structure and function based on quantitative expectations of what constitutes a fish community with high biotic integrity in a particular region and habitat type. All versions include metrics that address species richness and composition, indicator species, trophic function, reproductive function, and/or overall abundance and individual condition. Different metrics and metric expectations within each of these metric categories are what distinguish different IBI versions. A variety of approaches have been used to generate metric expectations, and the process of establishing appropriate, sensitive expectations is probably the most difficult step in preparing a new version of the IBI. At present, most existing IBI versions are for wadable warmwater streams in the central United States. However, versions have also been developed, or are in the process of being developed, for coldwater streams, large unwadable rivers, lakes, impoundments, and marine and Great Lakes estuaries in many different regions of the United States, and for streams and rivers in Canada, Mexico, France, and India. Despite the large amount of effort that has been directed towards IBI development, much remains to be done, both in terms of generating new versions for different regions and habitat types, and in terms of validating existing versions.

ACKNOWLEDGMENTS

The authors wish to thank the following individuals for providing information on research in progress: J.R. Karr, R.M. Hughes, D.L. Dycus, R.M. Goldstein, S.G. Paulsen, C. Yoder, R. Thoma, E.T. Rankin, E. Emory, J.R. Shulte, and G. Gibson. We appreciated critical review comments from three anonymous reviewers, which greatly improved an earlier draft of this manuscript. Support for John Lyons on this chapter was provided by the Wisconsin Department of Natural Resources, Bureau of Research and Bureau of Water Resources Management, Study RS60 FIS RS634.

REFERENCES

For references see:

Davis, W. S. and T. P. Simon (eds). 1995. Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making. (ISBN 0-87371-894-1) CRC Press: Boca Raton.