Development and application of a bioindicator for benthic habitat enhancement in the North Carolina Piedmont

Desiree D. Tullos\textsuperscript{a,}\textsuperscript{*}, David L. Penrose\textsuperscript{b}, Gregory D. Jennings\textsuperscript{c}

\textsuperscript{a} Oregon State University, Biological and Ecological Engineering Department, 116 Gilmore Hall, Corvallis, OR 97330, United States
\textsuperscript{b} North Carolina State University, Water Quality Group, Box 7637, Raleigh, NC 27659, United States
\textsuperscript{c} North Carolina State University, Biological and Agricultural Engineering, Box 7625, Raleigh, NC 27695, United States

\textbf{Article info}

\begin{itemize}
  \item Article history: Received 26 September 2005
  \item Received in revised form 2 March 2006
  \item Accepted 6 March 2006
\end{itemize}

\textbf{Keywords:}
Benthic macroinvertebrate
Habitat enhancement
Indicator species
Stream restoration
Reference condition
Bioassessment

\textbf{Abstract}

This paper describes the development, application, and evaluation of a method for assessing the effectiveness of stream restoration activities in enhancing four lotic habitats based on the presence of habitat specialists. Three genera were identified as specialists for indicators of the enhancement of woody debris, coarse bed substrate, fine roots, and leaf pack habitats. These indicator genera were determined for each habitat type through indicator species analysis, extensive literature review, and consultation with local experts and a statewide distribution database. Water quality influences were isolated by excluding taxa with low tolerance to degraded water quality conditions. The difference in the presence of indicator genera between pairs of upstream-restored reaches was used to evaluate the success of the restoration activities in re-establishing benthic habitats. Application of this methodology to 27 paired reaches in the North Carolina Piedmont indicated that no change in specialists was the most frequent result of restoration, particularly for the woody debris habitats, when each habitat was examined individually. By combining the habitats into a composite score, a distinction by land use emerged, with habitats in urban areas indicating the greatest enhancement, while presence of the indicator genera at the agricultural and rural sites showed no clear trend of improvement or degradation in response to the restoration activities. When this composite IG metric was compared to the EPT taxa richness metric and RBP scores, the dependency of the EPT taxa richness metric on upstream conditions and the improvement in discriminatory ability over the RBP score suggest that this indicator genera (IG) metric provides a distinct signal for representing the biological perspective on the enhancement of benthic habitats by stream restoration activities. While further development of the methodology is desirable, this framework introduces a valuable alternative for evaluating benthic habitat enhancement in various hydrogeographic and land use conditions, and is constructive for guiding restoration designs to maximize biotic integrity.

© 2006 Published by Elsevier B.V.

1. Introduction

Conversion of natural landscapes to urban and agricultural uses has resulted in the degradation of channel stability and aquatic communities. Channelization and the deforestation and farming of riparian areas are common practices in agricultural landscapes, while urban areas have converted forested river corridors to roadways and sewer right-of-ways, confined...
flow to the channel by altering and fixing channel boundaries and building levees, and altered hydrology of the watershed. Degradation of habitat that results from landscape development, including modifications to flow regimes, energy processing, and habitat access or quality, is considerably damaging to algal, fish, and insect communities (Paul and Meyer, 2001). As the natural flow regime is disrupted and riparian areas are converted to more anthropogenically productive uses, removal of woody debris, sedimentation of coarse bed substrates, decreased availability and retention of detrital material, and erosion of bank habitats occur. As the value of functioning and self-sustaining stream ecosystems becomes more widely recognized, attempts to mitigate the losses of these ecosystem functions have given rise to the science of stream restoration.

The National Research Council defined stream restoration as the “various techniques used to replicate the hydrological, morphological, and ecological features that have been lost in a stream due to urbanization, farming, or other disturbance” (National Research Council, 1992). Various practices have been applied in the attempt to restore these features through re-constructing natural pattern, profile, and dimension of the channel, re-establishing floodplain connection, and recreating accessible and high quality aquatic and terrestrial habitats. Because the hydrologic, morphologic, and ecologic features of rivers are intensely interrelated, successful restoration requires a comprehensive knowledge of the dominant influences on and the responses of the integrity of aquatic ecosystems (Nedeau et al., 2003). This is particularly true as restoration goals and criteria move beyond simply restoring channel form into restoring ecological functions (Palmer et al., 2005), necessitating the evaluation of ecological indicators to signal changes in stream ecosystem processes. While the ecology of rivers are dependent on the combined influence of the physical, chemical, and biological conditions of the river corridor, it is the physical habitat component upon which restoration activities often focus (Bernhardt et al., 2005).

The physical condition of a river reach is structured by both the local and larger scale influences, with aquatic habitat defined by instream features, watershed contributions of sediment and flow, and contiguous topography and activity. Researchers have reported the importance of high-quality habitats through the relationships between biological diversity and habitat quality (Raven et al., 1998), and the stress associated with impaired habitats (Karr et al., 1986; Karr, 1991; USEPA, 1977). Because diverse communities in aquatic ecosystems are largely dependent on functioning habitat (Barbour and Strickland, 1991; Plafkin et al., 1989; Southwood, 1977), re-establishing aquatic habitats have become a fundamental component of this developing science.

Thus, a key goal of stream restoration activities is often to re-establish high-quality aquatic habitats to facilitate the natural recovery process (Moerke et al., 2004). However, while it is widely recognized that diverse communities are consistently found in heterogeneous habitats (Pianka, 1967; Woodin, 1981; Boomsma and Van Loon, 1982; Schlosser, 1982), determining the success of re-establishing diverse and accessible habitats is rarely a trivial task. The measurement of habitat enhancement resulting from restoration activities is complicated by the lack of knowledge regarding (1) what level of biological variation naturally occurs following significant disturbances such as those created by constructing restoration activities (Nedeau et al., 2003), (2) the limitations on recovery imposed by activities in the watershed (Booth, 2005), (3) the timeline for recovery of biological communities following disturbance (Wallace, 1990), and the (4) ecological responses to physical restoration activities (Davis et al., 2003). However, in spite of these challenges facing biologically based assessment of habitat enhancement, determination of the success of restoration activities is imperative for those seeking mitigation credits (i.e., compensatory action to improve quality of a degraded stream for unavoidable impacts to another reach), as well as for documenting the effectiveness of measures for re-establishing functional stream ecosystems.

Two approaches have traditionally been applied to evaluate biotic integrity in aquatic ecosystems: enumeration of biotic communities into established metrics and qualitative or quantitative assessments of physical habitats. The use of benthic macroinvertebrates has been proposed for biological assessments of the habitat enhancement of reach-scale stream restoration (Brown, 2000), as they represent local conditions due to the restricted migration that characterizes many of these organisms (Potter et al., 2004), and play an important role in the food web of river systems (Covich et al., 1999). However, while it is recognized that benthic macroinvertebrates have distinct habitat preferences independent of water quality conditions, (Barbour et al., 1996), the direct use of benthic macroinvertebrates for detecting habitat enhancement by stream restoration activities is still under debate. This discussion is largely a result of the frequently inconsistent signals given by generic community descriptions and the inappropriate assumption that each metric signal similar ecological responses (Vlek et al., 2004). For example, Larsen et al. (2001) observed changes in the physical environment after adding woody debris to urban streams in the Puget Sound Lowland, but found no change in the biotic communities, as defined by the benthic index of biotic integrity (B-IBI) (Kerans and Karr, 1994). The authors also found that B-IBI values did not well correlate with channel characteristics, including woody debris, but correlated instead with level of urban development. Thus, it is unclear whether insects specializing in woody debris habitats did in fact return or were restricted by degraded watershed conditions. The use of stressor-specific biological metrics in assessing such stream rehabilitation projects may provide greater understanding of the mechanisms by which biotic integrity is achieved or limited.

While various metrics do exist for describing unique aspects of the macroinvertebrate community (Vlek et al., 2004), biotic indicators specific to habitat enhancement currently do not. Common multimetric indices, such as the B-IBI and the invertebrate community index (ICI) (Ohio EPA, 1987) are largely a function of diversity, which may not be sensitive enough to detect stressor-specific community shifts following habitat restoration. In fact, it has been found that IBIs were unsuccessful in expressing habitat degradation in northwest Mississippi (Shields et al., 1995). Alternatively, biotic indices,
such as the North Carolina biotic index (Lenat, 1993) and the family biotic index (Hilsenhoff, 1988), are based on organism’s tolerance to water quality conditions. However, the effects of stream restoration on water quality are largely undocumented (Jorgensen and Yarbrough, 2003) and because water quality serves such a distinct niche requirement from habitat (Barbour et al., 1996), the use of such indices is inappropriate for judging the success of restoration activities in enhancing habitats.

Ephemeroptera, Plecoptera, Trichoptera (EPT) taxa are often summarized into a commonly applied richness metric. EPT taxa richness represents a group of insects that are particularly sensitive to degraded water quality (Lenat and Penrose, 1996), and have been used to evaluate the success of stream restoration activities (Penrose, 2004). However, while the sensitivity of these organisms to poor habitat quality is recognized, this sensitivity varies across the orders (Karr and Chu, 1999), ranging from mayfly sensitivity to metal contamination (Kiffney and Clements, 1994; Yuan and Norton, 2003) to caddisfly dependence on availability of stable habitats (Death, 2003). Further, no explicit testing of the ability of the EPT taxa richness to reflect improvements in physical habitat initiated by stream restoration activities currently exists. Thus, it is uncertain whether these divergent sensitivities of the EPT orders create ambiguous signals regarding the success of stream restoration activities on improving benthic habitats since their responses are not specific to physical habitat quality. Because biological communities reflect the integrity of composite ecosystems, integrating physical, chemical, and biological features (Barbour et al., 2000), it is essential to define specific components of the community that reflect targeted responses if explicit evaluation of habitat enhancement is to occur.

An alternative to applying community data for the evaluation of physical habitat quality is the widely-used rapid bioassessment protocol (RBP), developed by Barbour et al. (1999). The RBP is a visually-based assessment of the level of impairment (or enhancement) of aquatic habitats. A set of 11 variables, reflecting habitat features defined by channel substrate and morphology, streambank characteristics, and riparian vegetation are assigned scores from 0 to 20, with higher values indicating higher quality habitats. These individual scores are then summed to reflect the overall diversity and quality of habitats in the stream reach. RBPs provide a quick and cost-effective evaluation regarding the quantity and quality of physical habitats, adopting the principle that these physical features are directly related to the quality and quantity of the biological community (Maddock, 1999; Rankin, 1995). However, as noted by Barbour et al. (1999), field crews must be trained before performing the RBP assessments, judgment criteria require calibration to the study area, and RBP scores involve periodic confirmation by local experts. Thus, to assure the quality and credibility of this semi-quantitative method, rigorous validation of the RBP scores is required to confirm the consistency and reliability of the assessments.

Thus, the applications of described traditional biotic indices and bioassessment techniques have conspicuous disadvantages in the assessment of habitat enhancement by river restoration practices. The isolation of water quality and habitat quality effects must be made to appropriately assess reach-scale restoration efforts without the burden of watershed-wide water quality obstacles to sensitive community re-establishment. For fair award of mitigation credits, variability and inconsistencies of field crews may be an unacceptable hazard and should be minimized in methods that rely on visually-based evaluations.

Thus, the development of new habitat criteria that distinguish meaningful biotic responses that are relevant to the stress of interest is clearly an important and urgent task (Ofenbock et al., 2004). This concern has led to the development and application of an alternative approach to assessing habitat enhancement as presented here, using the presence of indicator genera (IG) to signal the success of habitat enhancement. Specifically, the objectives of the study were to (1) develop the IG metric based on habitat specialists to target habitat enhancement signals, (2) apply the IG metric to 27 pairs of restored-upstream reaches in the North Carolina Piedmont to evaluate overall and individual habitat changes in response to restoration activities, (3) compare the IG metric with existing biologically and physically based indicators in the 27 paired sites, and (4) highlight the challenges and opportunities in developing biologically-based indicators of habitat enhancement.

2. Background

While the entire stream ecosystem provides habitat for the various benthic macroinvertebrates, this study focused on four habitat types in the North Carolina Piedmont streams: woody debris, coarse bed material, fine roots, and leaf packs. Two other significant habitat types not considered here include sand and aquatic mosses. Sand habitats were not considered because (1) the collection method (QUAL5) for this study does not sample low-velocity areas of streams, primarily pools, where sand is naturally found (NCDENR, 2003), and (2) the dominance of sand habitats in higher-velocity areas is not a desirable habitat to be ‘restored’ in these gravel-bed streams. Thus it is important to emphasize that the sites examined in this study are Piedmont streams; the application of this methodology in other ecoregions, such as the coastal plain, would require careful reconsideration of the natural and desirable habitat types and revised selection of indicator taxa.

The second habitat type not considered in this study is aquatic mosses. The role of mosses in stream ecosystem processes is significant (retention of fine particles, flow refugia for invertebrates) (Suren and Winterbourn, 1992). However, the incidence of aquatic mosses may be limited by their seasonality and moderate growth rate (Steinman and Boston, 1993). Further, it has been shown that mosses are often damaged during the construction of restoration activities by heavy equipment (Sand-Jensen et al., 1999). Because mosses are slow to mature and expand, their widespread presence following major disturbances, such as that associated with restoration activities, is unlikely (Muotka et al., 2002). These features of aquatic mosses make determining successful restoration of this habitat difficult and unreasonable within traditional monitoring periods.

Following is a detailed discussion of the four habitats considered for this study.
2.1. Woody debris

For the past two centuries, the removal of woody debris obstructions from streams and rivers was a prevalent management practice to enhance navigation and reduce flooding (Benke et al., 1985). Geomorphologic responses to woody debris removal include erosional downcutting (Bilby, 1984), widening (Maser et al., 1998), increased bedload transport, redistribution of gravel bars, and thalweg shifting (Smith et al., 1993a, b). Biologically, these woody ‘obstructions’ provide well-documented benthic habitat (Harmon et al., 1986) and have been shown by Benke et al. (1984) to comprise a significant fraction of the benthic macroinvertebrate production while only accounting for a minor proportion of the actual habitat. Further, the retention of organic material by woody debris plays a considerable role in the energy processing of the habitat. Further, the retention of organic material by woody debris to streams should result in more diverse and abundant benthic macroinvertebrate communities (Benke et al., 1985) and improve overall stream integrity.

2.2. Coarse bed substrate

Coarse bed substrates are a common and fundamental riffle habitat (Cummins and Lauff, 1969; Brown and Brussock, 1991) for the many clinger taxa, those organisms morphologically adapted to attaching to surfaces in riffles. It has been found that these habitats provide the greatest diversity and density of benthic macroinvertebrates (Hynes, 1970; Hart, 1978). However, these habitats are often and easily degraded in urban and agricultural landscapes when loads of fine sediments exceeding the transport capacity of the channel enter the stream from the landscape and by streambank erosion. These fine sediments embed larger particles by filling the spaces between them, removing access to interstitial habitats, and ultimately blanketing the coarse bed material that serves as critical habitat for many benthic macroinvertebrates. The effects of fine sedimentation are illustrated by the Stenonema mayfly larvae that are commonly found under loose substrates but rarely present in embedded substrates (Kondratieff and Voshell, 1980). Restoration of these coarse substrate habitats often requires re-establishing natural sediment transport and flow regimes, including minimization of excessive fine sediments from within and outside the channel.

2.3. Fine roots

Vegetation along streambanks provides necessary habitat for many of the climber insects, those that have evolutionary adaptations for vegetative habitats along streambanks, including overhanging branches and roots (Merritt and Cummins, 1996). Benthic larvae are often found in the vegetation of undercut banks (Bouchard, 2004) because adults, including the damselflies Argia and Calopteryx, will oviposit on the submerged vegetation (Borror, 1934). These rooty habitats are often removed by streambank erosion as shear stresses increase in response to disruptions in sediment transport and natural flow regimes. The restoration of such habitats is quickly accomplished, however, by re-establishing herbaceous and woody vegetation along streambanks (Manolis, 2003).

2.4. Leaf packs

Leaf packs provide another essential habitat for many benthic organisms, including the Peltoperlidae stoneflies, in addition to serving as the foundational layer of the food resources for the entire aquatic ecosystem (Vannote et al., 1980). While detrital material is often available in disturbed stream systems, the retention capacity of channels has often been lost due to stream management activities that removed retentive structures, such as woody debris and large bed material to increase navigation or protect property. Further, traditional management of urban and agricultural settings to increase productivity of an area has frequently resulted in the displacement and channelization of streams. In addition to the disruption of riffle–pool complexes, increased velocities, disrupted sediment transport, and destruction of habitat diversity (FISRWG, 1998), channelization also results in a substantial loss in the ability for a stream to retain allochthonous materials (Muotka et al., 2002). While supply is not always limited even in developed watersheds, the seasonality of detrital material from fall leaf litter (Cuffney and Wallace, 1989), makes retention a critical feature for sustaining benthic communities year-round (Lemly and Hilderbrand, 2000). However, the potential for restoration of these habitat types is great, since the increased presence of woody debris and large bed material, as well as increases in bed heterogeneity, has been shown to increase the retention of leaf material (Lemly and Hilderbrand, 2000; Negishi and Richardson, 2003; Benke et al., 1985).

Evaluating the success of restoration activities in re-establishing these four habitats is an important task for furthering the science of stream restoration and for awarding functionally-based mitigation credits. This paper describes the development, application, and evaluation of a method for assessing the effectiveness of stream restoration activities in enhancing these four habitats based on the presence of habitat specialists. First, a list of habitat specialists were defined as the IG list, with water quality sensitive taxa removed. For this methodology, the change in habitat quality from an upstream or baseline condition to a downstream restored site was used to evaluate the increase or decrease in presence of the IG in response to restoration activities. Using this difference approach, the IG metric was applied to analyze the improvement in habitat quality at 27 upstream-restored pairs in the North Carolina Piedmont. The results of these analyses were then compared to two commonly applied assessment methods: the EPT taxa richness metric and RBP scores. Finally, the challenges and opportunities of this metric are discussed.

3. Methods/materials

3.1. Study site characterization

A total of 27 pairs of upstream and restored reaches were studied across the Piedmont of North Carolina. Study
sites were sampled from May to August of 2003 and 2004 to minimize seasonal effects on environmental variables and aquatic insects. Sites were limited to drainage areas of 13 km² or smaller to further minimize environmental and community variation, with streams demonstrating a bankfull width range of 2.1–11.2 m. The soil conservation service curve number (SCS-CN) ranged from 54 to 83 (Soil Conservation Service, 1975), with landscapes varying from urban to agricultural to rural settings. Riparian area conditions included mature and immature forest, herbaceous cover, and lawn grass. Stream substrates ranged from silt to coarse gravel, with a median particle diameter range of 0.01–32 mm, and channel slopes from 0.07% to 2.67%.

Organisms were collected according to the North Carolina Division of Water Quality (NCDWQ) QUAL5 method (NCDENR, 2003). This protocol includes sampling insects from one riffle using a kicknet, one sweep net collection from bank habitats, fine mesh wash of rock/log, one leaf pack wash, and visual collections. Insects were picked from samples, preserved in 95% ethanol in the field, and brought back to the lab for identification to species when possible. Insects were not separated by habitat for these samples.

3.2. Study design

In an upstream–downstream approach, benthic macroinvertebrates were collected within each restored reach segment, as well as within a reach just upstream (100–500') of the project tie-in to investigate the effects of restoration activities on habitat quality given similar watershed conditions. It is important to note that several of the upstream reaches were not in what is often considered ‘reference’ condition. Particularly at a few of the urban sites, upstream conditions were notably degraded by watershed activities. However, in this modified upstream–downstream study design, the upstream reach was simply used as a control for comparison to the restored reach, to represent the baseline condition from which the restored reach was expected to deviate. This approach assumes that the watershed conditions similarly affect the upstream and restored reaches because baseline water quality and quantity conditions were similar to both reaches (Barbour et al., 1996), suggesting that differences in IG presences are a result of the restoration activities.

Further, it is worth noting that biological sampling occurred early in the recovery period for stream restoration, as the age of these projects ranged from 1 to 4 years. Given the limited understanding of the recovery time of restored streams, and that some of the study streams were potentially still stabilizing in response to construction activities at the time of assessment, this study does not attempt to investigate the temporal evolution of the sites, but instead takes a snapshot assessment of early responses of benthic macroinvertebrate communities to restoration activities. Thus, the term ‘restored’ within this paper refers to those sites to which stream restoration practices have been applied, without suggesting that sites have been fully restored to the natural, pre-development condition.

3.3. Indicator genera metric development

The enhancement of benthic habitat was investigated through the presence of indicator taxa, defined to be consistent and restricted to the group or habitat it represents (McCune and Grace, 2002). The implications are that the organism is always present and occurs primarily in a particular habitat, treatment, or area. Such biotic indicators can be used to identify ecosystem stresses due to their reliability to reflect current and cumulative ecosystem conditions (http://www.epa.gov/bioindicators/). For this study, three indicator genera were identified for each of the four habitat types through: (3.3.1) indicator species analysis, (3.3.2) a synthesis of habitat colonization studies, (3.3.3) removal of water quality sensitive taxa, (3.3.4) aggregation of species into genera, and (3.3.5) consultation with local experts on benthic ecology and query of NCDWQ basinwide monitoring database.

3.3.1. Indicator species analysis

The first task in developing the habitat specialist list was an indicator species analysis (Dufrene and Legendre, 1997) using data collected on Little Garvin Creek in the Piedmont of South Carolina (Young, 2004). The watershed for this stream is 6.8 km², with the channel characterized by an average water surface slope of 0.4% and bankfull width of 6 m. These features demonstrate the similarity in channel geometry at Little Garvin to the study sites evaluated within this paper, making the habitats and benthic organisms found in this reach appropriate for comparison with our study sites.

Benthic macroinvertebrates were sampled seven times in Little Garvin Creek. Each sample was collected and identified by habitat, resulting in a list of species found for each habitat type. These data were then used in an indicator species analysis within PCORD software (McCune and Mefford, 1999). This analysis combines the abundance and frequency of occurrence information of each taxon for each group into a single indicator value (IV). The groups analyzed for this analysis are the previously described four habitat types, with the faithfulness and abundance of each species collected at Little Garvin Creek providing the indicator value for each taxa in each habitat. One hundred Monte Carlo randomized simulations of the data were performed to ensure that these indicator values were statistically more significant than could occur by chance.

3.3.2. Synthesis of colonization studies

An extensive literature review followed, taking advantage of the numerous habitat colonization studies. These studies (Table 1) varied by ecoregion of study, stream type, and geographical location, but provided support for selected indicator taxa.

3.3.3. Removal of water quality sensitive taxa

Species with low tolerance values (TV), reflecting an organism’s sensitivity to water quality conditions, were then removed from the list. TVs applied in this study were developed through field tests of taxa abundance in five water quality categories (Lenat, 1993). The TVs, as listed in NCDENR’s ben-
Table 1 – Specialist list and supporting literature for each of the four habitat types

<table>
<thead>
<tr>
<th>Indicator genera</th>
<th>TV</th>
<th>Supporting literature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coarse bed material</td>
<td></td>
<td>Stenonema (Ephemeroptera) 5.5  Richardson and Tarter (1976), Kondratieff and Voshell (1980), Berner and Pescador (1988), Lamp and Britt (1981)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hydropsyche (Trichoptera) 7.8  Bouchard (2004), Georgian and Thorp (1992), Hickin (1968)</td>
</tr>
<tr>
<td>Fine roots</td>
<td></td>
<td>Calopteryx (Odonata) 7.8  Manolis (2003), Westfall and May (1996)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Argia (Odonata) 8.2  Borror (1934), Phillips (2003)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ischnura (Odonata) 9.5  Bouchard (2004), Cannings and Doerksen (1979)</td>
</tr>
<tr>
<td>Leaf packs</td>
<td></td>
<td>Tipula (Diptera) 7.3  Cummins et al. (1973), Peterson and Cummins (1974), Griffith and Perry (1993)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Leptophlebia (Ephemeroptera) 6.2  Edmunds et al. (1976), Peterson and Cummins (1974)</td>
</tr>
</tbody>
</table>

Tolerance values (TV) listed, with increasing tolerance values representing increasing tolerance to poor water quality, are for the most common species within the genus if not provided by genus in NCDENR, 2003. Literature supporting the use of indicator genera includes colonization studies as well as texts on the specific benthic organisms.

Thic Standard Operating Procedures (2003), range from 0 to 10, with 10 indicating high tolerance to pollutants. Thus, to isolate the effects of habitat enhancement from water quality effects, insects with a TV of 4.5 or lower were excluded from the set. Organisms below this tolerance value have demonstrated elevated intolerance to poor water quality conditions (NC DENR, personal communication). This exclusion of water quality influence was performed to confirm that changes in taxa presence reflected habitat restoration or loss rather than changes in water quality.

3.3.4. Aggregation of species to genera
Species were then combined into the genera taxonomic level for this list (Table 1). While it has been suggested that family-level taxonomy is sufficient for bioassessment programs (Hewlett, 2000), others have argued that species-level taxonomy is required (Resh and Unzicker, 1975; Bailey et al., 2001). In consideration of the evidence provided by these studies and due to the difficulty and reliability in identifying one of the selected indicator taxa (Stenochironomus) to the species level, we chose the genera taxonomic level for this study.

3.3.5. Consultation with local resources
The final list was used to query the NCDWQ basinwide monitoring database to confirm that the listed taxa do indeed occur in the river basins under study. This final step was taken to verify that aerial dispersal sources existed in the basin for re-colonization and that a lack of IG presence did not reflect inability to re-colonize. This indicator list was then confirmed by local experts with 20–30 years of field experience in the Piedmont of North Carolina to verify that insects were appropriately matched to habitats and representative of the basins under study.

3.4. Application of indicator genera metric
Each of the 27 pairs of upstream and restored study sites was queried to determine the presence of the identified habitat specialists. The presence of an indicator genus reflected one score point, so that sites with all genera for all habitats received a score of 12 (three genera per habitat, four habitats). The upstream score was then subtracted from the restored reach score to reflect the change in habitat quality from the baseline condition in response to restoration activities. With this approach, an increase in the presence of habitat specialists at a restored site, as compared to a paired upstream site, is taken to indicate an enhancement of habitat quality, with a decrease in specialists indicating the opposite. For example, a positive score on coarse bed material may suggest a success of the restoration activities in stabilizing streambanks and thus reducing fine sediment loads. The change in the presence of specialists was examined by each habitat type, as well as by the composite, overall habitat quality. An example demonstrates how these values were calculated.
Site D-upstream
Woody debris taxa: Macronychus (no) + Stenochironomus (no) + Ancyronyx (no) = 0 + 0 + 0 = 0
Coarse bed material taxa: Stenonema (no) + Stenacron (no) + Hydropsyche (no) = 0 + 0 + 0 = 0
Fine root taxa: Calopteryx (no) + Argia (no) + Ischnura (no) = 0 + 0 + 0 = 0
Leak pack taxa: Tipula (no) + Leptophlebia (no) + Isoperla (no) = 0 + 0 + 0 = 0
Upstream combined habitat score = 0

Site D-restored
Woody debris taxa: Macronychus (no) + Stenochironomus (no) + Ancyronyx (no) = 0 + 0 + 0 = 0
Coarse bed material taxa: Stenonema (yes) + Stenacron (no) + Hydropsyche (no) = 1 + 0 + 0 = 1
Fine root taxa: Calopteryx (no) + Argia (no) + Ischnura (yes) = 0 + 0 + 1 = 1
Leak pack taxa: Tipula (yes) + Leptophlebia (no) + Isoperla (no) = 1 + 0 + 0 = 1
Restored combined habitat score = 3
Composite site score = restored – upstream = 3 – 0 = 3

This example illustrates no improvement in woody debris habitat, but an improvement in the remaining three habitats for the restored reach is shown, resulting in an overall habitat score of +3 for the restored site.

3.5. Comparison of IG metric to EPT richness and RBP scores

EPT taxa richness and RBP scores were also calculated for the upstream and restored reach study sites. These two measures represent the information provided directly by the community in EPT taxa richness, and by visually-based habitat assessments with the RBP scores. RBP scores were calculated using the approach for high gradient streams due to the prevalence of riffle/run features in these Piedmont streams (Barbour et al., 1999). Evaluation of the three metrics were made through comparison of the coefficients of variation and simple linear regression. Finally, metrics were plotted by land use and time since construction to illustrate how these two influences were represented by each of the metrics.

The coefficient of variation (CV) is the ratio of the standard deviation to the mean, providing a relative measure of data dispersion compared to the mean when means between two datasets are unequal and when the units of measurement are different. The CV for each metric was compared to evaluate the reliability in discriminating between the upstream and restored reach sites. A substantially higher CV of one metric relative to the others indicates a lower reliability of the metric (Gomez and Gomez, 1976). For consistency, we compared sites based on the ratio of upstream to restored reach CVs so that higher ratios indicate higher reliability.

Simple linear regressions were performed for each metric between the upstream and restored reach sites to investigate the relationship of each metric to upstream watershed influences. High correlation between the upstream and restored reach values indicates a high dependence of the metric on the upstream watershed condition.

4. Results

4.1. Indicator genera metric development

The indicator species analysis proved to be minimally valuable. The limited robustness of the training dataset restricted generation of relationships between indicator species and associated habitats. While this analysis identified a few taxa

<table>
<thead>
<tr>
<th>Land use</th>
<th>Woody debris</th>
<th>Coarse substrates</th>
<th>Fine roots</th>
<th>Leaf packs</th>
<th>Overall site score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agricultural</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of positive</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>No. of negative</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>No. of no change</td>
<td>7</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Rural</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of positive</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>No. of negative</td>
<td>2</td>
<td>7</td>
<td>1</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>No. of no change</td>
<td>8</td>
<td>3</td>
<td>8</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Urban</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of positive</td>
<td>0</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>No. of negative</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>No. of no change</td>
<td>8</td>
<td>5</td>
<td>3</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>All land uses</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of positive</td>
<td>2</td>
<td>6</td>
<td>9</td>
<td>8</td>
<td>13</td>
</tr>
<tr>
<td>No. of negative</td>
<td>2</td>
<td>9</td>
<td>5</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>No. of no change</td>
<td>23</td>
<td>12</td>
<td>13</td>
<td>10</td>
<td>3</td>
</tr>
</tbody>
</table>

Each column represents the frequency of positive, negative, and zero differences in habitat genera specialists, calculated by subtracting the upstream, or baseline, condition from the restored reach score. The frequencies are compiled for all land uses at the bottom.
as reliable indicators of the various habitat types, only the Stenonema mayfly, an indicator for coarse bed substrate habitat, met the tolerance value criteria of greater than 4.5 at a statistical significance of \( p = 0.05 \). Consequently, the remaining habitat specialists (Table 1) were determined through colonization study reviews, with local expert consultations and NCDWQ database query verifying these taxa as appropriately matched to habitats and present in Piedmont stream systems.

4.2. Application of indicator genera metric

By examining the number of the positive, negative, and no-change values for each habitat individually, it appears that no change in specialists was the most frequent result at these sites (Table 2). The frequency of zeros indicates that neither improvement nor degradation to the quality of each habitat type can be attributed to the restoration activities at a majority of the sites. This trend was most consistent for the woody debris habitats, where 23 of the 27 sites indicated no change in specialists occurred. Further, the relative number of positive and negative composite habitat scores suggests that restoration activities had a negative effect on the presence of habitat specialists nearly as often as a positive effect. This is particularly true for the agricultural and rural projects, where the frequency of positively and negatively scored sites was similar. In contrast, the urban sites showed improvements in habitat, as indicated by the change in IG presence, at a majority of sites.

4.3. Comparison of IG metric to EPT richness and RBP score

By comparing the ratio of the coefficient of variation (CV) in the upstream to restored reaches, we evaluated the relative precision of the metric to discriminate between the upstream and restored sites, with higher ratios defining a higher relative precision or reliability. Keeping in mind that the upstream reaches represent the baseline condition of the streams prior to restoration activities, it appears that the IG metric has greater discriminatory power than the EPT richness and RBP scores, with CV ratios equaling 1.38, 1.14, and 0.82, respectively.

We also compared values in the upstream reaches to those in the restored reaches for each metric by simple linear regressions to evaluate the relationship between the restored and upstream reach scores. This was performed to investigate the independence of the metric on watershed-influenced, baseline conditions. Regressions performed on the three metrics indicated that the RBP was least dependent on upstream conditions, followed by the IG metric and EPT taxa richness, with \( R^2 \) values of 0.06, 0.17, and 0.71, respectively.

Finally, the difference in the scores of the upstream and restored reach metrics were plotted to visualize the similarity in response by land use (Fig. 1) and time since construction completion (Fig. 2). The land use plots demonstrate a similarity in the ability of the RBP and IG to detect a habitat quality change, particularly in the urban sites, whereas EPT taxa richness was unable to distinguish much change in the communities following construction. The plots by time since construction do not illustrate any clear difference in the sensitivity of the metrics with time, but do demonstrate that

---

**Fig. 1 – Effects of land use on the change in IG, EPT taxa richness, and RBP scores.** Again, the y-axis represent the restored reach value minus the upstream reach ‘baseline’ condition for the three metrics and the shaded line represents the mean value for each metric. These figures illustrate how the use of different metrics can lead to different conclusions regarding the benefit of restoration in the land uses. For example, RBP values suggest that agricultural watersheds appear to be more often degraded by restoration activities while EPT taxa richness indicate a more positive outcome at these same sites. In addition, the tightness of differences between upstream and restored reach differences indicates a decrease in the discriminatory ability of the metric, further illustrating the dependence of the metric on watershed conditions rather than between site differences.
slightly lower biotic metrics were measured at the older, more mature projects rather than the sites improving over time.

5. Discussion

The development and application of this methodology reflect an effort to improve restoration assessment and advance designs to re-establish aquatic habitats. While limited in its current application, this approach provides a promising framework for identifying relevant bioindicators and measuring their targeted responses to habitat improvements. Comparison of the IG metric with two common community assessment measures indicates that a unique signal is provided by this evaluation approach and that advantages do exist in using stressor-specific and biologically-based habitat assessments.

5.1. Application of the IG metric, individual habitats

A greater number of fine root habitats exhibited positive effects than any other habitat type (nine positive and five negative), which suggests that these habitats are more often successfully re-established than other habitat types. The restoration of bank habitats is promising as they both play a critical role as refugia in highly degraded systems (Roy et al., 2003) and serve as important habitats for dragonfly and damselfly larvae, which play a fundamental role as predators in the aquatic ecosystem (Strong and Robinson, 2004).

However, the parity of the positive and negative responses of habitat specialists to restoration for the remaining habitat types (woody debris: two positive and two negative; coarse substrate: seven positive and six negative; leaf packs: eight positive and nine negative) indicates that, generally, a majority of the study projects were unsuccessful in improving benthic habitat quality. The lack of response in habitat specialists was most notable for the woody debris habitat, with a strong majority of sites demonstrating no change in specialists for this habitat. Although re-establishing woody debris to streams has been identified as a measure for enhancing benthic productivity (Benke et al., 1985), these data illustrate little improvement in this critical habitat and suggest a need for focusing restoration designs to include woody debris as part of stream designs. Further, because woody debris is so often responsible for retention of detrital material, it is suggested that the return of woody debris to streams will positively influence the presence of leaf pack specialists as well. While it is discouraging to find that as many habitats appear to be degraded by restoration activities as are enhanced by them, it may reflect the design objectives during the early years of stream restoration science, and therefore be an inaccurate representation of current philosophies and design approaches.

5.2. Application of the IG metric, composite site scores by land use

By summing the four habitat scores into a single value, the net effect of the restoration activities was examined. The results for the urban sites are the most consistent and the most encouraging. While construction of dams and bridges, chan-
nelization, bank erosion, and loss of riparian habitats (Waters, 1995; Wood and Armitage, 1997, 1999) all result in degradation of benthic habitats in the urban watershed, it appears that stream restoration, to some extent, can overcome these negative impacts by enhancing local habitat conditions, illustrating the significant potential for re-establishing these habitats in urban areas. This is a unique finding and suggests that the barrier posed by water quality may obscure the effects of habitat enhancement in studies using untargeted ecological indicators. The agricultural and rural sites do not show the same clear overall improvement in habitat, with a similarity in the frequency of positive and negative scores. These findings suggest that these early restoration designs in agricultural and rural settings had not effectively enhanced benthic habitat at a majority of the sites at the time of assessment.

5.3. Comparison of IG metric with EPT taxa richness and RBP scores

Results from the metric comparisons indicate that the IG metric does provide a distinct response to habitat enhancement through the compilation of information directly from the biological community. The coefficient of variance ratios indicated a reduced precision of the semi-qualitative RBP scores, likely from the variability in field crew determinations in these visually-based assessments, justifying concerns that have been raised regarding the consistency of rapid habitat assessments (Stolnack et al., 2005). The linear regression of upstream to restored reach metrics suggests that the EPT taxa richness is strongly related to the upstream condition. This was interpreted as an indication of water quality influence on the EPT taxa and thus a lower sensitivity to habitat enhancement than the two habitat-targeted measures. Because the proposed IG metric is less sensitive to land use than the EPT taxa richness metric, and demonstrates a higher precision than the RBP scores, we believe this approach offers an improvement over EPT and RBP metrics because it is targeted to evaluate the underlying condition of interest (habitat enhancement) and is less sensitive to unrelated conditions (Patil, 1991), such as water quality.

The final comparison was to visually examine the similarities of the metrics in representing paired-site differences by land use and time since the completion of construction activities. These plots illustrate a decrease in variability of the scores in the EPT taxa richness, reinforcing the indication of a weakness in discriminating between sites with poorer water quality, as is often found in urban areas. It is suggested that these plots reflect the sensitivity of the EPT taxa richness metric to water quality, making its application in the evaluation of habitat enhancement questionable. The metric plots over time since construction show no clear trends in the metrics over time. However, both the IG metric and EPT taxa richness show a marginal decrease in the restored reach habitat quality at the more mature stream projects. Ostensibly, this finding contradicts the philosophy that stream recovery following construction is strengthened by time. However, we believe that these plots suggest how older projects lacked ecologically-based designs, and that, as the science of stream restoration matures, designs focused on channel stability have shifted to those with more ecosystem function objectives, including benthic habitat enhancement.

5.4. IG metric, challenges and opportunities to biologically based metrics

It has been recognized that the quality of aquatic habitats may be the most influential feature in structuring benthic communities (Rankin, 1995), and the application of stream restoration practices has great potential for re-establishing habitats in areas where they have been degraded. Understanding and measuring the effectiveness of restoration activities in enhancing benthic communities is important for determining whether projects meet stated goals, as well as for advancing the science of stream restoration. However, determining the success of restoration efforts is not a simple endeavor, and requires the selection of relevant criteria that reflect the transitions initiated by restoration activities (Brooks et al., 2002).

This paper describes a framework for evaluating the effectiveness of restoration activities on habitat enhancement. This evaluation procedure is independent of water quality conditions through the exclusion of water-quality sensitive taxa and provides an indication of the direction of change resulting from restoration activities without the requirement of determining a specific endpoint community. A central advantage of this approach is the lack of comparison between the ‘treatment’ stream with a ‘reference’ reach. Dominants-in-common (DIC) is one example of an approach that uses reference reach information, applied to determine the ecological condition based on the concept that dominant organisms represent the existing environmental features and influences (Shackleford, 1988). This metric requires the comparison of the treatment stream with a reference reach to determine the similarity of the dominant organisms, where the reference condition represents the least impaired streams (Parsons et al., 2002) and often the desirable endpoint community. However, this reference condition is often difficult to identify and impossible to restore due to irreconcilable watershed influences, particularly in urban stream systems.

Thus, in recognition of the complications associated with comparisons to the reference condition, the approach described in this paper evaluates an alternative perspective on restoration impacts by comparing post-restoration data to an upstream or pre-construction condition. The result is only the direction of the effect: Has habitat quality improved, decreased, or remained the same? This approach, while providing limited information on the magnitude of the impact, bypasses troubling questions:

- How much similarity to the reference condition is possible?
- How much similarity to the reference condition is enough to signify success?

Considering the limitations of stream restoration for solving watershed-wide development, the analyses described in this paper demonstrate the distinct and valuable contributions of targeted responses provided by indicator taxa. Particularly in urban areas where no appropriate ‘reference’ exists, the use of traditional water-quality sensitive criteria may not
be the best approach to evaluating the success of projects in enhancing benthic habitats.

It should be emphasized that the proposed IG metric is not intended to duplicate the assessment of water quality improvements. However, it could be combined into a multi-metric biotic index with other targeted metrics, such as water quality, temperature sensitivity, or nutrient processing, for developing a multi-component assessment of stream restoration projects. The biologically-based habitat assessment described within this paper is meant to reflect only one of the many interrelated components that support an aquatic ecosystem. The application of targeted metrics such as this is in the award of mitigation credits on a sliding scale by demonstration of recovery of specific ecological functions (habitat, water quality, flood retention, nutrient processing), and for evaluating the ability of various restoration practices to enhance specific ecosystem components.

5.4.1. Improvements and limitations

Jackson et al. (2000) proposed four phases for evaluating ecological indicators: conceptual foundation, feasibility of implementation, response variability, and interpretation and utility. While these four phases have been addressed in this paper, further study is required to adequately understand the significance of the limitations of this approach. These limitations address the second and third of Jackson et al.’s phases: the weaknesses of the indicator species analysis and the lack of documentation of sensitivity to natural and sampling variability. Weaknesses of indicator species analysis include, as with any bioindicator development, unidentified influences on indicator taxa and the requirement of a robust training dataset. First, unidentified influences may include changes initiated by the modification by restoration activities of the riparian area on shading, detrital inputs, and surface-water retention and treatment. Second, for this approach to be widely applied, robust training sets of benthic macroinvertebrates by habitat type should be used for indicator species analysis on an ecoregional basis.

Regarding the 3rd of Jackson et al.’s phases, LaPoint et al. (1996) noted the importance of natural variation and biases in the reliability of bioindicators. Though every attempt was made to isolate temporal and spatial variability, and while sampling for this study was performed by a consistent field crew, it is recognized that this could influence outcomes of any bioassessment, including that described here. Given the well-known relationship between sampling effort and the number of taxa collected (Larsen and Herlihy, 1998), it may be possible that this approach, like others, is susceptible to reflecting differences in sampling efforts rather than in actual differences in community. However, it is assumed that since the sampling method for the widely-used EPT taxa richness and the proposed IG metric is the same, errors and variability associated with collection are similar. Further, because this approach relies on only the presence, rather than abundance, of stressor-specific taxa that are known to exist in the river basins under study, we believe that the bias due to unequal sampling is low. Future research that includes a sensitivity analysis of the uncertainty in benthic macroinvertebrate metrics as a function of unequal sampling effort should indicate how the IG metric compares to existing metrics regarding this bias.

6. Conclusions

The development and application of this habitat restoration indicator provides insight into the measurement and determination of success in attempts to enhance a critical component of aquatic ecosystems. However, preserving natural habitats is fundamental to supporting robust benthic macroinvertebrate communities (Rankin, 1995), and the application of restoration activities indicates that an irretrievable loss has already occurred. Our success as engineers and ecologists in providing favorable conditions for these essential organisms may lie in our ability to protect and preserve the remaining high-quality streams, providing both habitat for the benthic organisms and opportunities for sustainable research by future students.

Acknowledgements

The authors would like to acknowledge the following agencies for financial support of this project: US Environmental Protection Agency (USEPA), North Carolina Department of the Environment and Natural Resources (NCDENR), Water Resources Research Institute (WRRI), USDA Cooperative State Research, Education, and Extension Service.

REFERENCES


Richardson, M., Tarter, D., 1976. Life histories of Stenonema cimicum (Walker) and S tripunctatum (Banks) in a West Virginia stream. Am. Midland Natural. 95, 1–9.


