River restoration, habitat heterogeneity and biodiversity: a failure of theory or practice?

MARGARET A. PALMER*,†, HOLLY L. MENNINGER‡ AND EMILY BERNHARDT§
*Chesapeake Biological Laboratory, University of Maryland Center for Environmental Science, Solomons, MD, U.S.A.
†Department of Entomology, University of Maryland, College Park, MD, U.S.A.
‡NY Invasive Species Research Institute, Cornell University, Ithaca, NY, U.S.A.
§Department of Biology, Duke University, Durham, NC, U.S.A.

SUMMARY

1. Stream ecosystems are increasingly impacted by multiple stressors that lead to a loss of sensitive species and an overall reduction in diversity. A dominant paradigm in ecological restoration is that increasing habitat heterogeneity (HH) promotes restoration of biodiversity. This paradigm is reflected in stream restoration projects through the common practice of re-configuring channels to add meanders and adding physical structures such as boulders and artificial riffles to restore biodiversity by enhancing structural heterogeneity.

2. To evaluate the validity of this paradigm, we completed an extensive evaluation of published studies that have quantitatively examined the reach-scale response of invertebrate species richness to restoration actions that increased channel complexity/HH. We also evaluated studies that used manipulative or correlative approaches to test for a relationship between physical heterogeneity and invertebrate diversity in streams that were not in need of restoration.

3. We found habitat and macroinvertebrate data for 78 independent stream or river restoration projects described by 18 different author groups in which invertebrate taxa richness data in response to the restoration treatment were available. Most projects were successful in enhancing physical HH; however, only two showed statistically significant increases in biodiversity rendering them more similar to reference reaches or sites.

4. Studies manipulating structural complexity in otherwise healthy streams were generally small in scale and less than half showed a significant positive relationship with invertebrate diversity. Only one-third of the studies that attempted to correlate biodiversity to existing levels of in-stream heterogeneity found a positive relationship.

5. Across all the studies we evaluated, there is no evidence that HH was the primary factor controlling stream invertebrate diversity, particularly in a restoration context. The findings indicate that physical heterogeneity should not be the driving force in selecting restoration approaches for most degraded waterways. Evidence suggests that much more must be done to restore streams impacted by multiple stressors than simply re-configuring channels and enhancing structural complexity with meanders, boulders, wood, or other structures.

6. Thematic implications: as integrators of all activities on the land, streams are sensitive to a host of stressors including impacts from urbanisation, agriculture, deforestation, invasive species, flow regulation, water extractions and mining. The impacts of these individually or in combination typically lead to a decrease in biodiversity because of reduced water.
quality, biologically unsuitable flow regimes, dispersal barriers, altered inputs of organic matter or sunlight, degraded habitat, etc. Despite the complexity of these stressors, a large number of stream restoration projects focus primarily on physical channel characteristics. We show that this is not a wise investment if ecological recovery is the goal. Managers should critically diagnose the stressors impacting an impaired stream and invest resources first in repairing those problems most likely to limit restoration.

Keywords: diversity, habitat, heterogeneity, invertebrate, restoration, river, stream

Introduction
Ecological science has a long history of using theory to guide and advance knowledge. Alongside prominent theories such as those related to island biogeography and trophic cascades, the role of habitat heterogeneity (HH) in promoting species diversity is one of the most often cited concepts in ecology (Ricklef & Schluter, 1993). Theoretical and empirical work on species diversity dates to the early days of ecological icons such as Hutchinson (1959) and MacArthur & MacArthur (1961), and there is a large body of research on the link between species diversity and HH or complexity (McCoy & Bell, 1991; Tews et al., 2004). Species diversity has been shown to increase with HH for a variety of species ranging from birds and mammals to insects and demersal fish (MacArthur & MacArthur, 1961; Murdoch, Peterson & Evans, 1972; Kaiser, Rogers & Ellis, 1999). The mechanisms are numerous and not necessarily mutually exclusive: HH may provide greater surface area, more physical refugia, and higher or more varied supplies of limiting resources. Thus, areas that in all other respects are ecologically equivalent but have a greater habitat variety may provide more ecological niches for members of a community, thereby promoting diversity (Warfe, Barmuta & Wotherspoon, 2008).

The assumption that high HH promotes biodiversity in stream ecosystems began appearing in the published literature over 30 years ago, and research directly comparing levels of species diversity across streams or stream reaches that differ with respect to geomorphic and/or in-stream HH began appearing regularly by the 1970s and continues today (e.g. Hynes, 1970; Allan, 1975; Williams, 1980; Muotka & Syrjanen, 2007; Jahnig, Lorenz & Hering, 2008). Stream ecologists recognised early on that the term heterogeneity was being used very loosely by ecologists (Erman & Erman, 1984) – it could refer to habitat complexity (technically the spatial arrangement of patches), habitat diversity (the number of types of habitats in an area), or even environmental variability within a habitat over time (Li & Reynolds, 1995). Attempts were made to go beyond correlative field studies in which it is difficult to separate heterogeneity from other confounding parameters that may also influence diversity such as average substrate particle size or food availability (Wise & Molles, 1979). Thus, a number of experimental studies began to emerge in which the diversity of particle sizes or their spatial arrangement was manipulated (e.g. O’Connor, 1991; Downes, Hindell & Bond, 2000). Despite mixed results from these studies, which we briefly review later, the assumption that HH promotes biodiversity in streams persists and has even been incorporated as a fundamental principle into textbooks (e.g. Allan & Castillo, 2007).

In addition to its theoretical basis, HH has received a great deal of attention in the environmental management arena because influencing habitat structure seems more tractable than influencing many other factors believed to support or enhance biodiversity (e.g. productivity, disturbance regime; Bell, Fonseca & Motten, 1997; Palmer, Hakenkamp & Nelson-Baker, 1997). Furthermore, arguments that species diversity may contribute to community stability and ecosystem function have increased the focus on biodiversity in the restoration context (Naem et al., 1994; Tilman, 1996). While the relationship between diversity and ecosystem function has been hotly debated, there is a general consensus that ecosystem function probably does decline as species are lost (Hooper et al., 2005; Cardinale et al., 2006). Indeed, a meta-analysis of work in eight different European grasslands suggests that different species have a disproportionate impact on different functions so that maintenance of multifunctional ecosystems may require maintenance of high species diversity (Hector & Bagchi, 2007). If both
functional diversity and response diversity within functional groups are high, an ecosystem may exhibit a great deal of resilience in the face of environmental changes (Elmqvist et al., 2003) while species poor assemblages may have a reduced capacity for community adaptation to compensate for changes in environmental conditions.

Given this theoretical backdrop, it is not surprising that today a central tenet of restoration ecology is that an ecosystem’s ability to withstand disturbances (i.e. be more stable) may be critical to its long term survival following restoration, and that this ability is enhanced when species diversity, and thus functional redundancy, is high (Falk, Palmer & Zedler, 2006; Lake, Bond & Reich, 2007). The desire to restore biodiversity in streams and rivers that have been degraded by land use change, agriculture, or other environmental stressors has primarily emerged over the last decade as the emphasis has shifted from restoration of single species (typically salmonid fisheries) to restoration of entire stream ecosystems and the suite of services they provide (Palmer et al., 2007). Along with this shift has come a strong focus on restoration of HH beyond the simple restoration of specific flows or habitats used by fish (Roni, Hanson & Beechie, 2008). Concordant with the shifted focus on HH was the increased usage of various measures of benthic invertebrate diversity, particularly of sensitive taxa, as indicators of stream health and restoration outcome (Karr & Chu, 1999; Allan & Castillo, 2007).

In 2000, field experiments were initiated to investigate the link between HH and restoration of biodiversity in streams. To our surprise, we found no change in invertebrate diversity in response to massive reach-scale manipulations of substrate heterogeneity (D84 : D50) as a restoration treatment for experimentally disturbed stream reaches (Brooks et al., 2002). At the time, we attributed this ‘negative’ finding to a lack of statistical power or an experimental scale that was inadequate. However, a few years later, Pretty et al. (2003) and Lepori et al. (2005) published complementary results indicating that increasing physical HH as part of restoration efforts on a number of European rivers did not result in increased biodiversity of fish or invertebrates. These papers were published at about the same time that the first comprehensive database on river and stream restoration was developed for the U.S. (Bernhardt et al., 2005; Palmer et al., 2005). This database showed that channel re-configuration and in-stream habitat improvements that increased heterogeneity were among the most common goals of U.S. stream restoration (Fig. 1). Similar databases were not available for other parts of the world, but literature from European projects clearly suggested a similar trend: river and stream restoration projects were relying heavily on channel re-configuration that added meanders and physical structures such as boulders and artificial riffles to enhance the structural heterogeneity of stream channels (Brookes, Knight & Shields, 1996; Harrison et al., 2004). As evidenced by interviews with project managers across the U.S. (Palmer et al., 2007), the focus on HH in river restoration is quite common and clearly reflects the assumption that habitat complexity promotes biotic recovery and particularly biodiversity (Harper et al., 1997; Palmer et al., 1997; Gerhard & Reich, 2000).

Because the prevalence of projects that manipulate stream channels remains high despite the costs (Bernhardt et al., 2007; Tullos et al., 2009), we felt it was time to revisit the links between geomorphic and in-stream HH and biological diversity in the context of restoration. Thus, we ask what evidence is there that stream restoration projects designed to enhance channel complexity or heterogeneity lead to an increase in aquatic biodiversity? To address this question, we synthesised the results of published studies that quantitatively evaluated invertebrate responses to the physical restoration of a degraded river or stream.

**Methods**

We conducted literature searches that included publications for the period 1975–2008. We did not restrict our searches to any particular journals, and we began by using the ISI Web of Science to search by crossing the following keywords: [Restoration OR Rehabilitation] AND [Stream OR River] AND Invertebrates. This search resulted in 113 articles (through December 2009). Each paper was scanned to assess its relevance to the motivating question and in a few cases authors were contacted to clarify issues and/or obtain original data. The criteria we used for relevance were: (i) the paper must describe results from monitoring the restoration of a degraded river or stream; (ii) the paper must provide quantitative data on invertebrate species richness; (iii) the paper must provide...
data on HH or else explicitly state that one of the restoration goals was to enhance habitat or channel complexity/heterogeneity; (iv) papers were not included that reported only on single habitats within a stream (e.g. invertebrate diversity on aquatic plants or wood with no information on the rest of the streambed) or that reported on a river or stream that was seriously polluted by mining or some other non-diffuse source. A number of papers were eliminated immediately based simply on the abstract; all other papers were read in full. We also evaluated the literature cited of every paper from the culled ISI search and found >30 potentially relevant articles that were not identified in the initial ISI search.

We extracted information from each paper on why, how and when the restoration was done, how channel or HH was assessed, the number of independent restoration projects that were evaluated, and the outcome of each project with respect to invertebrate species richness. In most cases, authors reported results of statistical analyses, and in all cases, we were careful to use the author’s written conclusions in evaluating project outcome with respect to invertebrate taxa richness (S = restoration significantly affected richness; NS = restoration had no significant impact; Inc = inconclusive results). Some of the papers reported on more than one project (e.g. Harrison et al., 2004; Tullos et al., 2009) but we only counted these as separate projects in our tally if they were truly independent of one another (i.e. in a different stream and part of a separate restoration project). In three cases, we had to use more than one publication by a group of authors to obtain all the needed information, and in two cases, original data on species richness were provided directly by the author (Muotka; Tullos). Our evaluation was meant to serve as a qualitative assessment of the effects of restoration of heterogeneity on biodiversity. A quantitative evaluation using some form of statistical meta-analysis to determine the magnitude of the response to heterogeneity was not possible because only a subset of the studies provided all the information required for such an analysis (e.g. sample sizes, variance).

Finally, to provide ancillary information, we completed the same literature search without the keyword ‘restoration’ and examined all papers that specifically tested for a relationship between physical heterogeneity in streams and invertebrate taxa richness (i.e. presumably in ‘healthy’ streams or rivers). The search produced 113 articles, a surprisingly similar number to the earlier search, but the articles did not overlap extensively. We also found a number of older papers that did not turn up in the ‘restoration’-excluded ISI search but were in the literature cited of some of the papers from that search. We used a similar process as described earlier to cull projects that were irrelevant; the criteria were the same except that no mention of restoration was necessary. We classified the relevant papers into two groups: studies that manipulated HH in an otherwise healthy stream and studies that attempted to correlate invertebrate diversity with

Fig. 1 Pie charts showing that channel re-configuration and in-stream habitat enhancements are among the most common restoration actions throughout the U.S. Data from Alexander & Allan (2007), Hassett, Palmer & Bernhardt (2007), Rumps et al. (2007) and Sudduth, Meyer & Bernhardt (2007).
existing HH. From each paper, we extracted information on the location of the study, what was measured, and the outcome.

Results

We found 78 independent restoration projects described by 18 different studies that met our criteria (Table 1). In all cases, efforts were made to enhance HH through restoration actions, and invertebrate taxa richness was measured in the restored reaches and/or at reference reaches upstream or in the watershed; some of the studies sampled pre- as well as post-restoration, but only a few did so at both restoration and reference sites (Table 1). Various anthropogenic stressors motivated the restoration efforts. Thirty-two of the projects were completed in the U.S. because of impacts from urbanisation or agriculture, while the remaining projects were performed in European, Australian, or Asian rivers to correct past practices of channelisation. Most of the projects involved geomorphic reconfiguration in the form of altering the channel, re-grading banks, creating more pool-riffle complexes, and adding in-stream structures such as boulders and wood (Table 1). In some cases, HH (also termed complexity) was quantitatively evaluated by measuring attributes such as reach-scale variance in particle size or current velocity in the restored versus unrestored reaches (as in Gerhard & Reich, 2000; Negishi & Richardson, 2003; Harrison et al., 2004; Lepori et al., 2005; Jähnig, Lorenz & Hering, 2009); however, in many cases only qualitative statements were made regarding enhanced heterogeneity as a desired project goal or outcome (as in Biggs et al., 1998; Friberg et al., 1998; Larson, Booth & Morley, 2001).

In three of the studies, biological monitoring was completed <2 years following the restoration treatment (Tikkanen et al., 1994; Biggs et al., 1998; Nakano & Nakamura, 2006; Walther & Whiles, 2008); however, some of the studies monitored projects that had been completed up to 16 years previously (Laasonen, Muotka & Kivijärvi, 1998). The study that examined the largest number of independent projects (in 24 streams), completed monitoring 2–10 years post-restoration with only one sampling visit per project (Tullos et al., 2009). Half of the studies involved monitoring invertebrates in the same stream multiple years following restoration (Table 1).

Of the 18 studies, two reported on one project in which the investigators measured significant increases in taxa richness post-restoration (Edwards et al., 1984; Jungwirth, Moog & Muhar, 1993); one study (Gerhard & Reich, 2000) reporting on two restoration projects found higher taxa richness in one of these projects, while the other showed no significant change in response to the restoration. We report the results of Nakano & Nakamura (2008) as inconclusive because taxa richness increased within one sub-habitat type of the restored reach (‘edge habitats’) but was not significantly higher at the reach-scale 15–27 months following restoration and their papers make different statements about the restoration outcomes (footnote 3, Table 1). The most comprehensive study in terms of number of projects monitored, methodological rigour and statistical analysis (Tullos, Penrose & Jennings, 2006; Tullos et al., 2009) reported no significant effect in sixteen restoration projects in agricultural or rural areas; for eight urban projects, they report a slight increase in taxa richness in restored reaches but largely the same taxonomic composition as in unrestored reaches.

The remaining 13 studies representing 66 separate restoration projects reported no significant change in taxa richness in the restored versus reference reaches (Table 1). In every case, we examined findings for the last post-restoration date provided since biotic response could take time and projects varied greatly in terms of how long after restoration they were sampled (1–16 years).

Our search also found 15 studies that manipulated HH in otherwise ‘healthy’ streams (i.e. streams not slated for restoration) and measured biodiversity following the manipulation (Table 2). Six of these studies reporting on nine independent manipulations found no significant effect of their manipulation on invertebrate species richness, seven found significant results and one was inconclusive. For all except one study that involved whole-riffle manipulations (Brooks et al., 2002), the spatial scale at which the biotic response was measured was 1 m² or less; most studies measured invertebrate diversity on ceramic tiles, blocks, colonisation trays or small patches of streambed in which habitat complexity had been manipulated using wood, stones, grooves, or artificially created roughness.

Twelve comparative survey studies measured ambient HH and invertebrate species richness in
Table 1 Restoration of heterogeneity and stream invertebrate diversity. Summary of publications reporting on restoration projects with a goal of enhancing in-stream habitat heterogeneity (HH); number in parentheses indicates how many independent restoration projects were evaluated. Effect on taxa richness in restored versus reference reaches was either significant (S; with \(P\)-value if provided), non significant (NS), or analysis of the publication(s) was inconclusive (Inc). In all cases, results are at reach or larger scales. In many cases, HH was not quantified using a metric but was described in the article or photographic evidence was available. Text in quotes reflects authors’ statements.

<table>
<thead>
<tr>
<th>Study</th>
<th>Location (# projects)</th>
<th>Land use</th>
<th>Restoration and monitoring approaches</th>
<th>Significance of effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Edwards et al. (1984)</td>
<td>OH, U.S.A. (1)</td>
<td>Hardwood stands in floodplains, row-crops in watershed</td>
<td>Artificial riffles and pools created by layering boulders over earthen fill in a previously channelised river. Post-restoration monitoring on restored, unrestored and least-impacted ('natural') reaches conducted in the 4th, 5th and 6th year post-construction</td>
<td>S – (P &lt; 0.001); Indicates project was meant to enhance HH by providing different depths and velocities. Reports that the number of invertebrate families was higher in 'natural' (17.7) and restored (15.1) versus unrestored (11)</td>
</tr>
<tr>
<td>Jungwirth et al. (1993)</td>
<td>Lower Austria (1)</td>
<td>Pastureland</td>
<td>Channel re-configuration completed to remove ‘pavement’, widen and add more habitat features to a channelised river; pre- and post-monitoring each of 3 years after project</td>
<td>S – Reports wider range of substrate types and increased spatial variance in depths and velocities. Reports an increase in the number of species between restored and channelised reaches but no raw data nor statistics</td>
</tr>
<tr>
<td>Tikkanen et al. (1994)</td>
<td>Northern Finland (1)</td>
<td>Coniferous forest with open canopy</td>
<td>Bulldozer ‘reworked beds’ to create more heterogeneity and boulders added to channel; upstream comparison site with pre- and post-monitoring five times for up to 30 days post-restoration</td>
<td>NS – Reports ‘slight increase in bed roughness &amp; mean particle size with a trend towards more turbulent, hydraulically rough conditions’. Reports no ‘measureable effect’ on species composition; standard error bars in figures but statistics not reported</td>
</tr>
<tr>
<td>Biggs et al. (1998)</td>
<td>Jutland, Denmark (1)</td>
<td>Pastureland</td>
<td>Reconfiguration to add meanders, allow floodplain overflow, added gravel and cobble, created habitats with backwater eddies and pools; monitored pre- and 1 year post-restoration in restored and upstream reaches</td>
<td>NS – Photographs show dramatic increase in channel meandering and in-stream substrate heterogeneity. For both rivers, the macroinvertebrates pre versus post did not differ in number of taxa nor compared to upstream reach</td>
</tr>
<tr>
<td>Friberg et al. (1998)</td>
<td>Jutland, Denmark (1)</td>
<td>Unknown, likely farmland</td>
<td>Reconfiguration of channelised reach to add 16 meanders, added gravel and rock material; pre- and post-monitoring in restored and upstream reaches; post-monitoring done four times until 6th year post-restoration</td>
<td>NS – ‘Restored reaches were significantly more heterogeneous’ than upstream reaches. Taxa richness relative to upstream reach fluctuated from year to year; lower in 4th year but higher in 6th year; ‘overall diversity [of upstream reach] similar to that of restored reach’. No statistics reported</td>
</tr>
<tr>
<td>Laasonen et al. (1998)*</td>
<td>Northeastern Finland (6)</td>
<td>Mostly forested</td>
<td>Previously channelised streams restored to enhance ‘habitat heterogeneity’; monitoring completed 8 or 16 years post-restoration in restored and reference sites</td>
<td>NS – Relative bed roughness higher in restored channels and additional depths and flows were present compared to unrestored. Species richness not higher (in some cases lower) than in channelised (unrestored) reaches</td>
</tr>
</tbody>
</table>
Table 1 Continued

<table>
<thead>
<tr>
<th>Study</th>
<th>Location (# projects)</th>
<th>Land use</th>
<th>Restoration and monitoring approaches</th>
<th>Significance of effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gerhard &amp; Reich (2000)</td>
<td>Hesse, Germany (2)</td>
<td>Spruce, alder,</td>
<td>Channelised sections received inputs of large wood – the Joseklein River received them ‘passively’</td>
<td>$S$ (Joseklein), $NS$ (Lude) – Quantified number of microhabitats/reach including woody debris, boulders, ‘stones’, sand and leaves; restored reaches had more microhabitat patches per metre than unrestored reach. Species richness higher in restored reaches of Joseklein ($P$-values not reported), but no difference between restored and unrestored in the Lude</td>
</tr>
<tr>
<td></td>
<td></td>
<td>beech forest</td>
<td>following tree cutting and floods; the Lude had ‘bundles’ added. Post-restoration monitoring</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>completed 4 years in both restored and unrestored reaches</td>
<td></td>
</tr>
<tr>
<td>Larson et al. (2001)</td>
<td>WA, U.S.A. (4)</td>
<td>Urban</td>
<td>Degraded urban streams rehabilitated to ‘enhance habitat’ by adding wood that resulted in</td>
<td>$NS$ – ‘Significantly increased channel complexity’; No significant difference in invertebrates between manipulated and reference sites (used an Index of Biotic Integrity which included data on species richness)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>‘geomorphic changes to channel’; Pre-restoration samples collected for three streams; post-restoration monitoring for 2–10 years in restored and upstream sections</td>
<td></td>
</tr>
<tr>
<td>Muotka et al. (2002)†</td>
<td>Northern Finland (7)</td>
<td>Deciduous forests</td>
<td>Formerly channelised reaches restored ‘using earth-moving machinery’; improved heterogeneity by placing boulder dams, deflectors, cobble, pebbles and wood in channel; Pre- and post-restoration monitoring 2, 4, 8 years after restoration in restored, unrestored and least disturbed reference sites</td>
<td>$NS$ – Restoration resulted in a statistically significant increase in channel complexity. Authors suggest channels are on a ‘trajectory to recovery’ based on invertebrate abundance and some taxa but indicates they do not have all the taxa that natural streams have</td>
</tr>
<tr>
<td>Muotka &amp; Syrjanen (2007)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Purcell et al. (2002)</td>
<td>CA, U.S.A. (1)</td>
<td>Urban, suburban</td>
<td>Removed culverts and re-configured channel to desired sinuosity; added step pools, rocks and riparian vegetation to highly degraded urban stream. Monitoring completed 3 years post-restoration in restored and unrestored reaches; also monitored project restored 10 years ago as a comparison</td>
<td>$NS$ – Increased channel complexity by meanders, step pools but also increased vegetation in buffers and removed all concrete. Taxa richness in 1999 overlapped in all three sites: restored (22.7 taxa), previously restored (22.3) and unrestored (22); no statistical analysis provided; Oddly, abstract and discussion conclude ‘biological condition improved’</td>
</tr>
<tr>
<td>Negishi &amp; Richardson (2003)</td>
<td>BC, Canada (1)</td>
<td>Coniferous forests</td>
<td>Placed boulder clusters in a stream reach that had been ‘anthropogenically altered’ (simplified streambed) due to bridge construction; no excavation activities involved. Sampled treatment reach and two upstream reaches twice pre- and once post-restoration in year 1, three times in year 2</td>
<td>$NS$ – ‘Increased habitat heterogeneity’ compared to pre-treatment and compared to reference reach; assessed statistically using coefficient of variation of velocity, substrate and depth. Restoration had ‘no significant impact on taxa richness’; ‘effect negligible’. Statistical analysis completed but $P$-levels not provided</td>
</tr>
<tr>
<td>Study</td>
<td>Location (# projects)</td>
<td>Land use</td>
<td>Restoration and monitoring approaches</td>
<td>Significance of effects</td>
</tr>
<tr>
<td>----------------------</td>
<td>-----------------------</td>
<td>---------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Harrison et al.</td>
<td>Lowlands, U.K. (13)</td>
<td>Agricultural</td>
<td>Restored channelised streams by adding in-stream flow deflectors made of large cobble and/or adding gravel and cobble to create heterogeneous riffles; Monitored ‘rehabilitated’ and nearby ‘un-rehabilitated reference’ in each of 2 years that were 4–9 years after rehabilitation</td>
<td>NS - ‘Increased flow and depth heterogeneity’ relative to control reaches (coefficient of variability in flow and depth reported in Fig. 5 of Pretty et al., 2003). No statistically significant difference in ‘taxa richness between rehabilitated stretch and reference stretch’; $P = 0.19$ and 0.23. (Note – author indicates substantial difference in species richness between habitats within reaches; complex paper due to design)</td>
</tr>
<tr>
<td>Moerke et al.</td>
<td>IN, U.S.A. (1)</td>
<td>Riparian trees and shrubs intact, formerly agriculture</td>
<td>Channel reconfiguration for habitat improvement on a channelised stream; created ‘natural’ pool-riffle sequences, added wood, boulders and gravel; stabilised streambanks, planted vegetation. Monitored restored and un-restored reference reaches once pre-restoration and once per year for 5 years post-restoration</td>
<td>NS - Restored reaches with significantly higher habitat quality [used an index that includes diverse types of habitats]. No significant difference in Shannon Diversity between restored and un-restored reaches for any of the 5 years post-restoration</td>
</tr>
<tr>
<td>Lepori et al.</td>
<td>Northern Sweden (4)</td>
<td>Managed forest</td>
<td>Reconfigured to remove bank armoring, widened channel, added new habitat features including boulders to a previously channelised river; post-restoration monitoring on restored and two un-restored streams; restorations occurred at different times 3–8 years prior to study</td>
<td>NS - Heterogeneity quantified by direct measures of bed profiles; statistically higher heterogeneity in restored versus unrestored sites (heterogeneity was 8 SD greater in restored reaches versus unrestored). No statistical difference in species richness between restored and unrestored ($P = 0.769$)</td>
</tr>
<tr>
<td>Nakano &amp; Nakamura</td>
<td>Northern Japan (1)</td>
<td>Photo in 2008 paper shows pastureland and forest</td>
<td>Channel reconfiguration in 2002 to re-create meandering in one reach and add rock groynes in another reach that had been channelised; post-restoration sampling in each of two habitat types (edge, mid-channel) of each reach completed five times (3, 8, 15, 20, 27 months)</td>
<td>Inc – Re-meandered reach had significantly more edge habitat which had different depths, velocities and sediments; reach with groynes with more variable flows. Taxa richness in edge habitat of restored reach higher than in edges of channelised reach but in mid-channel habitat no difference between restored and unrestored reaches. Statistical analysis indicates taxa richness in restored reach after 15, 20 and 27 months was not different from (unrestored) reach</td>
</tr>
<tr>
<td>Jähnig &amp; Lorenz</td>
<td>Central highlands, Germany (7)</td>
<td>Forest (60%) and pasture + agriculture (30%)</td>
<td>Three previously channelised reaches ‘actively’ restored to create a multi-channel (braided) form by removing bank ‘reinforcements’ and widening channels while four were restored ‘passively’ by not maintaining channelisation and allowing multiple-channel formation. Post-restoration sampling completed once for each site 2–9 years following restoration</td>
<td>NS – Extensive measurements of habitat complexity including coefficient of variation in depths and velocity; also report number of habitat types, spatial arrangement and substrate size. Restoration resulted in statistically significant increase in heterogeneity. Author concludes: ‘Despite higher habitat diversity, alpha [taxa] diversity has not changed’</td>
</tr>
<tr>
<td>Study</td>
<td>Location (# projects)</td>
<td>Land use</td>
<td>Restoration and monitoring approaches</td>
<td>Significance of effects</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-----------------------</td>
<td>-----------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>Tullos et al. (2009)</td>
<td>NC, U.S.A. (24)</td>
<td>8 urban, 8 rural and 8 agricultural</td>
<td>Evaluated eight projects in which restoration involved reconfiguring channels in streams that had ‘lost channel complexity’ (based on morphometric measures and amount of large wood debris/length of stream). Monitoring conducted 1–4 years post-restoration; 24 streams not all restored at the same time.</td>
<td>NS (rural, agricultural) – 16; S/NS (urban) – 8; For rural and agricultural catchments, average channel complexity lower in restored reaches; in urban areas not significantly different. Taxa richness averaged over eight sites not significantly different for rural and agricultural streams but taxonomic composition was. In urban streams, average taxa richness across eight sites was slightly higher in restored reaches but species composition no different – both had depauperate fauna. Authors conclude: ‘restoration acted as a disturbance, a filter against organisms’ [of certain functional groups]</td>
</tr>
<tr>
<td>Walther &amp; Whiles (2008)</td>
<td>IL, U.S.A. (1)</td>
<td>Forest, agriculture and wetlands</td>
<td>Constructed riffles and pools by adding 25 rock weirs (‘constructed riffles’) to a channel that had been modified, had flow diversions and was geomorphically degraded. Post-restoration sampling completed 1 year following implementation of projects in restored and unrestored reaches. NS – In-stream ‘hydraulic heterogeneity’ increased as well as presence of rock structures and new riffle/pool complexes. Raw data for taxa richness not provided but authors report: ‘Mean taxa richness across all sampled habitats was 22.0 ± 1.2 (mean ± 1 SE) in 2003 and 23.8 ± 1.1 in 2004. Rock weirs did not have higher richness and did not result in higher diversity at the reach level despite differences in composition’</td>
<td></td>
</tr>
</tbody>
</table>

*Authors reported on 13 streams but only the six streams that had been restored 8 or 16 years prior to monitoring are independent of the streams studied in Muotka et al. (2002).†Muotka supplied raw data to confirm our understanding of the published work.

‡Results from Tables 1 & 2 in Nakano & Nakamura (2008) show taxa richness initially increased (3 and 8 months post-restoration) but later was significantly lower than reference reaches (15, 20, 27 months post-restoration); however, in a different publication (Nakano et al., 2008) authors stated that ‘taxa richness was always higher [in the restored reach] than the channelised reach’. 
Table 2  Studies manipulating habitat heterogeneity (HH). Studies reporting on projects in which invertebrate species richness was tracked after enhancing physical HH in a stream not designated for restoration. All projects occurred in streams with forested buffers except three: Stewart et al. (2003), Arrington et al. (2005) and Schneider & Winemiller (2008). When significance levels for statistical tests were provided, they are reported along with the conclusion of S, NS, or Inc (significant, non-significant, inconclusive, respectively)

<table>
<thead>
<tr>
<th>Study</th>
<th>Location and land use</th>
<th>Heterogeneity manipulation</th>
<th>Significance of effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allan (1975)</td>
<td>CO, U.S.A.</td>
<td>Mixed stream substrate of different sizes in colonisation pans placed in riffles</td>
<td><strong>NS</strong> – Concluded that sampling and colonisation trays with mixed substrates ‘suggestive’ of a heterogeneity effect but no consistent results</td>
</tr>
<tr>
<td>Wise &amp; Molles (1979)</td>
<td>NM, U.S.A.</td>
<td>Mixed stream substrates of different sizes in colonisation baskets put in riffles</td>
<td><strong>NS</strong> – Concluded that increasing substrate heterogeneity did not increase species richness</td>
</tr>
<tr>
<td>Williams (1980)</td>
<td>ON, Canada</td>
<td>Mixed four sizes of stream substrates in wire baskets and placed in riffles</td>
<td><strong>NS</strong> – Concluded that substrate size, not heterogeneity, was driving species diversity and composition.</td>
</tr>
<tr>
<td>Erman &amp; Erman (1984)</td>
<td>Three independent experiments: two in CA, U.S.A. One in CO, U.S.A.</td>
<td>California: mixed stream substrates of different sizes but held median size constant</td>
<td><strong>NS</strong> – California streams, <em>P &gt; 0.05</em></td>
</tr>
<tr>
<td>Clifford et al. (1989)</td>
<td>AB, Canada</td>
<td>Rough versus smooth tiles</td>
<td><strong>S – <em>P &lt; 0.05</em></strong></td>
</tr>
<tr>
<td>O’Connor (1991)</td>
<td>Vic., Australia</td>
<td>Blocks of wood with and without grooves</td>
<td><strong>S – <em>P &lt; 0.0001</em></strong>, higher species richness on grooved wood</td>
</tr>
<tr>
<td>Douglas &amp; Lake (1994)</td>
<td>Melbourne, Australia</td>
<td>Bricks, grooved in various ways</td>
<td><strong>S – <em>P &lt; 0.0001</em></strong>, species richness increased with surface complexity</td>
</tr>
<tr>
<td>Death, R. (2006), unpubl. data</td>
<td>North Island, New Zealand</td>
<td>Bricks, grooved or not</td>
<td><strong>NS</strong></td>
</tr>
<tr>
<td>Downes et al. (1998)</td>
<td>Vic. Australia</td>
<td>Bricks with added pits and cracks</td>
<td><strong>NS/S – <em>P = 0.094</em></strong> on day 14; <em>P = 0.048</em>** on day 28</td>
</tr>
<tr>
<td>Downes (2000)</td>
<td>Vic., Australia</td>
<td>Roughened versus smooth bricks</td>
<td><strong>S – <em>P &lt; 0.001</em></strong></td>
</tr>
<tr>
<td>Brooks et al. (2002)</td>
<td>VA, U.S.A.</td>
<td>Manipulated substrate size in 12 reaches of a stream so that <em>D_84 : D_50</em> was high, medium, or low but median size held constant</td>
<td><strong>NS – <em>P = 0.314</em></strong></td>
</tr>
<tr>
<td>Stewart et al. (2003)</td>
<td>VA, U.S.A.</td>
<td>Attached varying numbers of stones to concrete slabs and placed these in the river</td>
<td><strong>S – <em>P &lt; 0.006</em></strong></td>
</tr>
<tr>
<td>Taniguchi &amp; Tokeshi (2004)</td>
<td>Reihoku-Machi, Japan</td>
<td>Stone plates with variable numbers of cavities</td>
<td><strong>S – In three of four seasons, <em>P &lt; 0.05</em></strong></td>
</tr>
<tr>
<td>Arrington et al. (2005)</td>
<td>Plains region, Venezuela</td>
<td>High versus low complexity patches created using ceramic blocks</td>
<td>**Inc – Significance unresolved; invertebrate results not analysed separately in the paper (contacted author who believes there was a difference)</td>
</tr>
<tr>
<td>Schneider &amp; Winemiller (2008)</td>
<td>TX, U.S.A.</td>
<td>High versus low complexity patches created using sticks</td>
<td><strong>S – Higher family richness on more complex patches</strong></td>
</tr>
</tbody>
</table>

streams that were mostly forested (many in national parks or reserves) (Table 3). Of these, four found a statistically significant relationship, while eight failed to find such a relationship.

**Discussion**

Regardless of the type of stressors that have impacted a stream or river, a focus on physical stream characte-
Table 3 Diversity–heterogeneity surveys. Studies in the literature in which physical habitat heterogeneity (HH) and invertebrate species richness or diversity were measured in a healthy stream; dominant watershed land use is indicated if reported by authors. Scale for sampling and evaluating heterogeneity (riffle, reach, etc.) indicated. When significance levels for statistical tests were provided, they are reported along with the conclusion of S or NS (significant, non-significant, respectively).

<table>
<thead>
<tr>
<th>Study</th>
<th>Location</th>
<th>Scale (# streams evaluated)</th>
<th>Land use</th>
<th>Heterogeneity measure</th>
<th>Significance of heterogeneity effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beisel et al. (1998)</td>
<td>Northern France</td>
<td>Reach and sub-reach (four streams)</td>
<td>Forest</td>
<td>Patch diversity and within – patch diversity (substrate size, flow, depth)</td>
<td>NS – No difference in species richness across patches that differed in substrate or patch ‘richness’ (i.e. heterogeneity) but abstract makes conflicting statements S – Within-patch (bryophyte patches) complexity was associated with more species</td>
</tr>
<tr>
<td>Harper et al. (1997)</td>
<td>Ireland; Czech Republic</td>
<td>Reach (multiple ‘rivers’)</td>
<td>Forest</td>
<td>Habitat functional diversity (# of patch categories)</td>
<td>NS – Could not distinguish between main channel (with lower habitat diversity) and floodplain regions with high habitat diversity; attribute it to water quality differences</td>
</tr>
<tr>
<td>Minshall &amp; Robinson (1998)</td>
<td>ID, U.S.A.</td>
<td>Reach (32 streams)</td>
<td>Forest</td>
<td>Coefficient of variation in substrate size</td>
<td>NS – Compared substrate Coefficient of Variability to species richness among the three groups of streams</td>
</tr>
<tr>
<td>Robson &amp; Chester (1999)</td>
<td>Hobart, Tasmania</td>
<td>Riffles (one stream)</td>
<td>Unknown land use</td>
<td>Fractal dimension within patches and number of patches within a riffle (substrate size)</td>
<td>S – P &lt; 0.05; More species on boulder-cobble riffles than in bedrock riffles</td>
</tr>
<tr>
<td>Buffagni et al. (2000)</td>
<td>North Italy</td>
<td>Reach (one river) ‘unaltered with high water quality’</td>
<td></td>
<td>Identified dominant habitats (units) and characterised each based on substrate, flow, depth, roughness</td>
<td>NS – Paper focused on showing that species habitat type (unit) is not a surrogate for biota but authors also provide data on roughness, means and SD for flow, depth, grain size. The number of taxa in a habitat unit was not clearly related to the Coefficient of Variability of flow, depth and grain size or roughness</td>
</tr>
<tr>
<td>Brown (2003)</td>
<td>NH, U.S.A.</td>
<td>Riffles (one stream)</td>
<td>Forest</td>
<td>Substrate heterogeneity (size) using diversity indices</td>
<td>S – Biodiversity increased and temporal variability decreased with heterogeneity</td>
</tr>
<tr>
<td>Boyero &amp; Bosch (2004)</td>
<td>Panama</td>
<td>Riffles (one stream)</td>
<td>Forest</td>
<td>Variability in substrate types (qualitatively assessed size) and velocity in riffle</td>
<td>S – Species richness was statistically higher on stones in riffles that had varying substrate sizes and variation in velocity across the riffle; potential problem is authors only sampled three stones per riffle</td>
</tr>
<tr>
<td>Urban et al. (2006)</td>
<td>CT, U.S.A.</td>
<td>Multi-scale study: riffle to watershed (18 streams) Streams along urbanisation gradient</td>
<td></td>
<td>Variation in habitat, substrate and flow</td>
<td>NS – Most important variable explaining diversity was riparian vegetation and landscape scale (watershed) factors</td>
</tr>
</tbody>
</table>
characteristics persists as a driving force in the motivation and design of many restoration projects – almost half of the individual projects we found were reported on in 2006–08 alone. Diverse reasons were given for the restorations ranging from past channelisation, various urban impacts, agricultural run-off, flow diversion, or culvert placement; yet, all restoration projects involved some form of channel reconfiguration and enhancement of in-stream structural complexity (Fig. 2). Across all of the studies we evaluated, only three reported significant increases in species richness and one was inconclusive (Table 1). Further, one of these studies found this result for only eight of the 24 restoration projects they evaluated and all eight of those were in urban streams in which there was only a marginal increase in the number of taxa, all of which were typical opportunistic, urban taxa (Tullos et al., 2009). The taxonomic composition was no different from what was found in the unrestored ‘control’ reaches despite the extensive nature of the restoration projects; the authors concluded that restoration was minimally effective in urban settings (Tullos et al., 2006). This means that across the 78 independent restoration projects monitored by the 18 sets of studies we evaluated, only two of the 78 projects resulted in increases in invertebrate diversity sufficient for the authors to conclude that the project was a biological success.

What can explain this finding? One might conjecture that increases in diversity in response to restoration focused on the channel and habitat characteristics were not found because the studies were not sufficiently rigorous in a design or statistical sense. This explanation is particularly appealing given the increased emphasis on effectiveness monitoring in the last 5 years and the number of publications on monitoring practices (e.g. Bernhardt et al., 2005; Falk et al., 2006). However, the best example of a rigorous study of those we reviewed (Tullos et al., 2009) reported results that are very much in line with our conclusions; these authors used replication, had control sites, and used multiple metrics to evaluate biotic outcome yet still found that the majority of projects did not result in increased species diversity.

Perhaps then, the lack of response to heterogeneity was simply because restoration project evaluations were conducted without allowing sufficient time for biological recovery; however, many of the studies sampled multiple years following restoration including several for 8–10 years. It is also not possible to
uniformly blame the lack of response to a lack of real increases in HH because many of the studies quantified an increase in heterogeneity. Lepori et al. (2005) reported an eightfold increase in heterogeneity in the restored versus unrestored reaches, and the projects evaluated by Harrison et al. (2004) had significantly higher heterogeneity (quantified with coefficient of variation) in flow and depth (Pretty et al., 2003). However, at least one study reported that increases in HH did not necessarily persist a long time even though restoration designs suggested they should; many variables describing habitat complexity and stability did not differ between restored and unrestored sites (Tullos et al., 2009). This is an interesting result in itself because it suggests that aside from biological outcome, restoration practices may have limited effectiveness geomorphically in some settings. In previous evaluations of in-stream enhancement projects, Roni et al. (2005) have also emphasised that the durability of in-stream enhancements is highly variable.

Perhaps, invertebrates do not respond to the types or spatial scale of habitat enhancements (Brooks et al., 2002), while other aquatic biota do. In-stream habitat enhancement projects are very common for freshwater fish and some studies have shown that salmonids respond to additions of in-stream structures (largely wood) particularly if the restoration includes reconnection of critical habitats (Roni et al., 2008). However, Steward et al. (2009) formally synthesised empirical evidence for the effectiveness of in-stream structural additions in restoring salmonids and concluded that the widespread use of these structures is not supported by the data. Structures may attract fish (aggregate them) but this does not mean fish communities are restored.

Is it possible that ecological theory on HH and biodiversity is simply lacking and should not be guiding restoration efforts? To address the basic theory aspect of this question, we went back to the literature and evaluated research that tested for a relationship between HH and invertebrate diversity in healthy streams, i.e. streams not in need of restoration. We found that about half of the studies (9/16) that experimentally manipulated heterogeneity in streams found statistically significant evidence that taxa richness increased (Table 2). For studies that asked if natural variation in HH explained variation in invertebrate diversity, about a third (4/12) reports a significant relationship (Table 3). Thus, despite strong evidence from many terrestrial systems and some aquatic systems that more heterogeneous habitats support higher levels of biodiversity (MacArthur & MacArthur, 1961; McCoy & Bell, 1991; Kerr & Packer, 1997; Tews et al., 2004), the evidence that geomorphic and in-stream HH strongly influence invertebrate diversity in healthy streams remains an open question.

With respect to the question of whether channel alterations and habitat enhancements should be the primary factor guiding stream restoration efforts, our findings clearly indicated it should not if the goal is ecological recovery. Our results show that for the majority of restoration studies to date, biodiversity was
not increased following attempts to enhance heterogeneity, and, only a subset of studies in healthy streams have demonstrated a positive biotic response to high levels of channel or in-stream heterogeneity. We do not entirely dismiss the idea that heterogeneity may play an important ecological role in streams but suggest that because so many other factors influence stream biodiversity, focusing solely on habitat structure as the basis of restoration efforts makes little sense. Heterogeneity may act in concert with factors such as disturbance regime, food resources and regional species pools (Lake, 2000; McCabe & Gotelli, 2000; Ward & Tockner, 2001) to influence diversity (Menninger & Palmer, 2006; Muotka & Syrjanen, 2007; Warfe et al., 2008), but there is little evidence that habitat heterogeneity is a primary determinant of macroinvertebrate diversity. In fact, for many restoration sites, water quality may not be sufficient to restore invertebrate diversity even if heterogeneity were restored – i.e. heterogeneity may not have a ‘chance’ to play an important role in promoting diversity when other constraints such as high pollutant loads, degraded hydrological regimes, or lack of a colonist pool exist.

We did not set out to identify which of the many factors may have influenced restoration success or lack thereof in the studies we evaluated – the information in the studies is not sufficient to undertake such a task. However, from attempts by others to identify the most important factors influencing stream invertebrate diversity in watersheds with varying levels of impact, it is clear that HH does not rank highest on the list. Urban et al. (2006) report that habitat was not as important a predictor as riparian vegetation and watershed landscape structure, while Kiffney & Roni (2007) point towards the importance of food availability. In general, studies show that the diversity and composition of biotic communities in streams depend strongly on factors at multiple scales with catchment-scale variables such as land use being most important (Townsend et al., 2003; Walsh, Fletcher & Ladson, 2005). Roni et al. (2008) emphasise the important role of larger-scale factors in promoting restoration success and advise the following sequence of actions for rehabilitating streams and rivers: protect critical habitats, improve water quality, restore watershed processes (e.g. habitat connectivity, hydrology), and then improve in-stream habitat. This suggests that if water quality, flow and riparian conditions are adequate then biota may indeed respond to heterogeneity. Thus, first and foremost, restoration efforts must target the most limiting factor, i.e. the factor that must be corrected before biota can return. Most often that factor is altered water quantity and quality and these can only modestly be improved with local (reach-scale) interventions, if at all. Practitioners recognise this constraint and water quality improvement is the primary goal of many river restoration projects (Bernhardt et al., 2007), yet practitioners assume that structural improvements will promote water quality – an assumption even more tenuous than the link between HH and macroinvertebrate diversity.

Given the large number of projects that rely on reach-scale, structural approaches, our findings have serious implications for future restoration efforts. Much more must be done to restore streams than simply re-configuring channels and enhancing structural complexity with meanders, boulders, wood, or other structures. The studies we evaluated suggest little to no improvement or perhaps even a decline in stream health because the restoration actions involved in these projects may act as a disturbance (Tullos et al., 2009). Further, research that is ongoing in our own labs indicates that restoration involving traditional channel manipulations often result in a decline in invertebrate and plant diversity. If these restoration approaches are not working, this begs the question: what are the alternatives? First, ‘softer’ restoration approaches that do not involve full-scale manipulation of a channel or its riparian corridor are prudent. Second, larger-scale actions such as storm-water management, changes in forestry or agricultural practices, and preservation of land and riparian vegetation may be the best hope for mitigating the impacts of anthropogenic activities in streams (Larson et al., 2001; Walsh, Fletcher & Ladson, 2005; Muotka & Syrjanen, 2007; Saunders & Fausch, 2007).

Given the lack of evidence to date that the costs of channel re-configuration and addition of in-stream structures (Bernhardt et al., 2005) are offset by biological recovery, we suggest that resources should be put into other actions (Lave, Robertson & Doyle, 2008). Not only are watershed-scale actions known to lead to significant improvements in the ecological health of streams, but whole-watershed perspectives can provide insight into the probability of success of a specific project in a specific place. The extent and position of degraded land, the spatial arrangement of tributaries and the connectivity of suitable habitat all act to
determine restoration outcome at individual reaches (Palmer, 2009). While we are far from identifying the relative importance each of multiple factors plays in limiting the success of restoration, we can apply a logical, data-driven approach to choose among various restoration methods and to prioritise sites for restoration. Data such as those presented in the papers we evaluated are critical to making progress in restoring our waterways.

Acknowledgements

This paper was presented at the September 2008 multiple stressors symposium in Windermere, U.K. We wish to thank: the sponsors and organisers of the meeting, most particularly the Freshwater Biological Association of the U.K.; an anonymous reviewer, and constructive comments from J.D. Allan. This work was supported by a grant from the David and Lucille Packard Foundation as well as support from the EPA Collaborative Network for Sustainability (Award # 8832206). This is contribution number 4314 from the Collaborative Network for Sustainability (Award # 8832206). This is contribution number 4314 from the University of Maryland Center for Environmental Science.

Conflicts of interest

The authors have declared no conflicts of interest.

References


© 2010 Blackwell Publishing Ltd, *Freshwater Biology, 55* (Suppl. 1), 205–222


(Manuscript accepted 16 November 2009)