

Review

Evaluating Stream Restoration Projects: What Do We Learn from Monitoring?

Zan Rubin ¹, G. Mathias Kondolf ^{1,*} and Blanca Rios-Touma ^{1,2}

¹ Department of Landscape Architecture and Environmental Planning, University of California Berkeley, Berkeley, CA 94720, USA; zankrubin@gmail.com (Z.R.); briostouma@gmail.com (B.R.-T.)

² Facultad de Ingenierías y Ciencias Agropecuarias, Ingeniería Ambiental, Unidad de Investigación en Biotecnología y Medio Ambiente -BIOMA-, Campus Queri, Calle José Queri, Edificio #8, PB, Universidad de las Américas, 170504 Quito, Ecuador

* Correspondence: kondolf@berkeley.edu; Tel.: +1-510-664-7804

Academic Editor: John S. Schwartz

Received: 1 September 2016; Accepted: 7 February 2017; Published: 28 February 2017

Abstract: Two decades since calls for stream restoration projects to be scientifically assessed, most projects are still unevaluated, and conducted evaluations yield ambiguous results. Even after these decades of investigation, do we know how to define and measure success? We systematically reviewed 26 studies of stream restoration projects that used macroinvertebrate indicators to assess the success of habitat heterogeneity restoration projects. All 26 studies were previously included in two meta-analyses that sought to assess whether restoration programs were succeeding. By contrast, our review focuses on the evaluations themselves, and asks what exactly we are measuring and learning from these evaluations. All 26 studies used taxonomic diversity, richness, or abundance of invertebrates as biological measures of success, but none presented explicit arguments why those metrics were relevant measures of success for the restoration projects. Although changes in biodiversity may reflect overall ecological condition at the regional or global scale, in the context of reach-scale habitat restoration, more abundance and diversity may not necessarily be better. While all 26 studies sought to evaluate the biotic response to habitat heterogeneity enhancement projects, about half of the studies (46%) explicitly measured habitat alteration, and 31% used visual estimates of grain size or subjectively judged ‘habitat quality’ from protocols ill-suited for the purpose. Although the goal of all 26 projects was to increase habitat heterogeneity, 31% of the studies either sampled only riffles or did not specify the habitats sampled. One-third of the studies (35%) used reference ecosystems to define target conditions. After 20 years of stream restoration evaluation, more work remains for the restoration community to identify appropriate measures of success and to coordinate monitoring so that evaluations are at a scale capable of detecting ecosystem change.

Keywords: stream restoration; evaluating success; post-project appraisal; metrics; aquatic macroinvertebrates

1. Introduction

With increasing popularity of stream restoration in the US, a number of publications in the early 1990s argued for more monitoring and evaluation of projects, so that the experience gained from current projects could be used to improve future endeavors (e.g., [1–4]). These calls for more stringent evaluations of restoration success presented an opportunity to treat restoration actions as experiments to develop better understanding of the river systems and test out approaches. The need for monitoring and evaluation was echoed in subsequent works, such as post-project appraisal approaches proposed by Downs and Kondolf [5], and in detailed guidance for selecting metrics and indicators for restoration offered by Woolsey et al. [6]. Palmer et al. [7] identified “pre- and post-assessment”

and public availability of data as one of five criteria for successful projects. Evaluation has become increasingly common, though many evaluations do not definitively answer if restoration has succeeded. Many metrics have been used, but they are not always tied to project objectives, nor necessarily appropriate for measuring the changes effected by the restoration interventions. Perhaps the most universal insight from multiple evaluations of stream restoration is the importance of understanding the complexity of stream systems and their potential responses to restoration.

1.1. Challenges to Evaluation

It may be easier to call for evaluation of restoration projects than to actually carry it out. As Rutherford et al. [8] warned, “it is often unwise for managers to evaluate the bio-physical impacts of their interventions unless they do it ‘properly’”. Routine or casual monitoring is unlikely to demonstrate change resulting from a small-scale restoration action for multiple reasons: the restoration effect (even if successful) may be too small to produce a measurable result, the restoration project may involve multiple actions whose effects may be difficult to distinguish, baseline data may be inadequate, or the measured variables may have high natural variability [8]. Using data for the Latrobe River, Australia, Rutherford et al. [8] demonstrated that to detect a statistically significant decrease in turbidity of 10% would require 80 years of sampling. Larger reductions could require less time to document. In some cases, a properly designed and executed evaluation study could be more expensive than the restoration action itself, a reality that is difficult for many managers to accept (or at least to sell to the public). The investment needed for meaningful assessment implies that restoration evaluation may be possible only in some cases, and effective evaluation strategies may require pooling resources across multiple projects.

1.2. Shortcomings of Commonly Used Evaluation Metrics

Many project proponents, funders, and (for mitigation projects) regulatory agencies have adopted linear or areal measures, such as area of riparian habitat or length of stream restored. For example, to comply with requirements from the Office of Management and Budget to quantify outputs from its ecosystem restoration projects, the US Army Corps (the Corps) reports “acres restored” or “acres restored per \$1 spent” [9]. Areal metrics do not distinguish high quality habitat (or habitat critically needed for a given species and life stage) from lower quality habitat (or habitat that was not critically needed by important species). The Corps has also undertaken efforts to develop more sophisticated metrics of restoration, developed to evaluate the likely benefits of ecosystem restoration and tradeoffs between economic and environmental benefits [10–13].

Monitoring target populations may also be ineffective as an evaluation metric. Throughout northwestern North America, habitat restoration projects are implemented to benefit anadromous salmonids, but the populations of these fish are notoriously influenced by other factors such as climate and ocean conditions. Indeed, after 80 years of constructing in-stream structures for salmonid habitat, there is no clearly documented evidence of population increases as a result [14] a sign of the challenges in using population monitoring to evaluate restoration success. As noted by Krebs [15], “Monitoring of populations is politically attractive but ecologically banal unless it is coupled with experimental work to understand the mechanisms behind system changes”.

1.3. Biological Integrity Indices

One of the main applications of biological indices is as an indicator of water quality, the idea being that while periodic sampling of water chemistry could easily miss a transient pulse of pollution, the macroinvertebrate community would reflect the water quality over a long time period, with more pollution-sensitive taxa absent in streams subject to poor water quality. However, in the context of restoration, many projects have not intended to improve water quality per se, but to improve instream habitat through additions of rocks and logs, or even through complete channel reconstruction. It is not immediately clear that indices based on pollution sensitivity of macroinvertebrate taxa are appropriate

to measure effectiveness of restoration projects, though they are commonly used. Habitat heterogeneity projects may also be seeking to reduce streambank erosion and fine sediment is a common water-quality impairment. In such cases, use of a biological index may be warranted. However, as reviewers of these studies, we were surprised that authors frequently did not justify the appropriateness of their chosen metrics.

1.4. The Opportunity to Look Back

Although rigorous (multiple years of pre- and post-project evaluation, quantitative measurements of habitat conditions, etc.) monitoring of restoration projects remains the exception rather than the rule [16], over the past two decades there have been many careful studies of the performance of restoration projects, and these provide us with a potentially important database with which to assess the effectiveness of different metrics to evaluate success. Two relatively recent papers by Palmer et al. and Miller et al. independently reviewed multiple studies of restoration projects, all of which were intended to enhance habitat heterogeneity (i.e., complexity of pool-riffle structures, undercut banks, large wood, etc.) [17,18]. While the meta-analyses of Palmer et al. and Miller et al. sought to assess whether restoration programs were succeeding, our review focuses on the evaluations themselves, and asks what exactly we are measuring and learning. We were intrigued by these two reviews (cited more than 400 times, with individual studies collectively cited 1500 times), and the fact that they reached divergent conclusions. We carefully analyzed all the original studies used in the two reviews. We sought to determine how the different methods and metrics used might have influenced determination of “success”. We also critiqued the underlying premise that macroinvertebrate diversity and richness should be universally applicable metrics of restoration success. Although we focus our review on the evaluation of habitat heterogeneity enhancement projects, (e.g., remeandering, rock and wood structures, etc.) we draw upon examples of other restoration approaches and expect that the principles and considerations will be applicable to the evaluation of many types of restoration projects.

2. Review of Habitat Heterogeneity Enhancement Restoration Evaluation Studies

In our systematic review of methods and metrics in the 26 studies used by Palmer et al. [17] and Miller et al. [18] (Table 1), we categorized each study according to twelve criteria representing a range of important considerations for restoration evaluation as presented in the literature (e.g., [7,19]). We documented location, extent of pre- and post-project monitoring (to control for temporal variability); sampling frequency (to control for seasonal variability); and whether the studies sampled different habitats (e.g., pool, riffle, banks) separately, in aggregate, or were restricted to certain habitat types (Table 2). We identified underlying assumptions and evaluation methods employed in each study, and considered these in light of observational and theoretical studies of food webs and species interactions. We also noted whether control (degraded, unrestored) sites were included to control for temporal variability and provide basis for comparison, whether regional reference (e.g., nearby, best potential ecological condition, *sensu* Reynoldson et al. [20]) sites were included to control for spatial variability and provide a regionally appropriate standard of success, what standards of success were stated and employed, and whether studies measured habitat heterogeneity directly, visually, or not at all. Other considerations were whether potential construction impacts from the restoration project were considered, and whether the regional reference sites had comparable drainage areas and slopes (Table 2). We scrutinized the methods and designs of each individual study, to assess the likelihood that the study designs would be adequate to detect biological change.

Table 1. Attributes of Studies Reviewed by Miller et al. (2010) and Palmer et al. (2010).

Study	Location	Habitat Metrics	Biological Metrics
Biggs et al. 1998 [21]	Brede river, Denmark; Cole river UK		Taxonomic richness, abundance/cover and species rarity (data available for the UK only).
Ebrahimnezhad & Harper 1997 [22]	Harper's Brook, England	D, V	Diversity and taxonomical differences among sites.
Edwards et al. 1984 [23]	Olentangy River, OH, USA		Taxonomic richness, abundance, diversity, biomass. For fish (number, biomass, species).
Friberg et al. 1998 [24]	Jutland, Denmark	D, S, V	Taxonomic Richness, density, composition. Abundance of stone-dwelling species were analyzed to specific level.
Gerhard & Reich 2000 [25]	Central Germany	D, F, L, S, V, W	Maximum and average number of species and density.
Gortz 1998 [26]	River Esrom, Denmark	C, O, S, V	Saprobic index, Danish fauna index, invertebrate diversity, count of trout spawning redds.
Harrison et al. 2004 [27]	Several locations in UK	D, S, V	Taxonomic richness, Shannon diversity (H'), total abundance and the abundances of individual taxa.
Jähning & Lorenz, 2008/Jähning et al 2009/Jähning et al. 2008 [28–30]	Lahn, Eder, Nims and Bröl Streams in Germany.	B, D, S, V, W	Taxonomical richness, similarity on composition among substrates and channel types (single or multiple).
Jungwirth et al. 1993 [31]	Epipotamal and Melk Rivers, Austria	D, S, V	For macroinvertebrates number of species and drifting biomass. For fish number of species and diversity.
Laasonen et al. 1998 [32]	Iijoki, Oulankajoki and Oulujoki in Finland	D, S, V	Taxonomical richness, density and assemblage structure. Shredders density.
Larson et al. 2001 [33]	Pudget Sound Tributaries, WA, USA	D, L, P, S	Index of Biological Integrity (IBI).
Lemly and Hilderbrand 2000 [34]	Stony Creek, VA, USA	P	Taxonomical richness and functional feeding groups (FFG).
Lepori et al. 2005 [35]	Ume River, Sweden	D, M, Q, S, V, W	Taxonomic richness, composition, evenness of fish and macroinvertebrate. Taxonomical density for macroinvertebrates.
Lester et al. 2007 [36]	Gippslandand, Australia	W	Number of families, abundance, number of EPT families, average SIGNAL2 sensitivity score, Shannon's Evenness index (Shannon's E), and FFG.
Moerke et al. 2004 [37]	Juday Creek, IN, USA	C, G, L, S	Macroinvertebrate density and diversity. Benthic algae density. Fish richness and biomass.
Muotka & Laasonen 2002 [38]	Northeastern Finland	D, M, O, S, V	Densities of macroinvertebrates and FFG.
Muotka et al. 2002 [39]	Same as Laasonen et al 1998	D, S, V	Densities of macroinvertebrates and FFG.
Nakano and Nakamura 2006 [40]	Shibetsu River, Northern Japan	D, S, V, Y	Taxonomic richness, density and composition.
Negishi & Richardson 2003 [41]	Southwestern British Columbia, Canada	D, S, V	Macroinvertebrate abundance, rarefacted richness and community composition.
Pedersen et al. 2007 [42]	Skjern River, Denmark	D, S, V, W	Macrophyte number of species and coverage. Macroinvertebrates: richness, abundance, Shannon diversity, EPT and dominant species frequency.
Purcell et al. 2002 [43]	S.F Bay, CA, USA	G	Taxonomic richness, family richness, total number of individuals, Family biotic index, EPT richness, % EPT individuals.
Roni et al. 2006 [44]	Umpqua and Coquille basins, OR, USA	L, P, S	Total abundance, Taxonomical richness, relative abundance (proportion of total abundance) of FFG (shredders and collectors) orders EPT tax , and I-IBI.
Rosi-Marshall et al. 2006 [45]	Cook's Run, MI, USA	D, G, O	Periphyton concentration and biomass. Macroinvertebrates: density, diversity and FFG composition. Fish: Trout abundance, % Age 1+Trout; % harvestable trout.
Sarriquet et al. 2007 [46]	Tamoute River, France	H, Q, S	Composition, density, taxonomical richness.
Tullos et al. 2009 [47]	North Carolina Piedmont, USA	E, G, O, P, S	Taxonomical and Trait composition, and Shannon Diversity. Species indicator analysis of restored and un restored sites.
Walther & Whiles, 2008 [48]	Cache River, IL, USA		Taxonomical diversity, richness, evenness and similarity, composition of assemblages. FFG.

Notes: B is braiding index, C is canopy cover or shading, D is depth, E is bank erosion, F is facies mapping, G is generic quality assessment (EPA's RBP, or Ohio's QHEI), H is hyporheic exchange, L is count or volume of wood, M is moss cover, O is Organic Content of substrate or leaf retention, P is pool spacing or pool area, Q is water quality, S is substrate size, V is velocity, W is width, Y is shear stress, FFG is functional feeding groups.

Table 2. Sampling design summary.

Location	14 (54%) of projects were in Europe. 10 (38%) were in North America, 1 (4%) in Australia, and 1 (4%) in Asia
Pre-project monitoring	16 (62%) had no pre-project monitoring. 9 (35%) monitored for one year and 1 (4%) monitored two years
Post-project monitoring	16 (62%) had one year of post-project monitoring. 5 (19%), 3 (12%), 1 (4%), 1(4%) sampled 2,3,4, and 5 years respectively
Sampling frequency	14 (54%) sampled once per year. 7 (27%), 4 (15%), and 1 (4%) sampled 2, 3, and 4 times per year respectively
Habitat sampled	5 (19%) sampled a single habitat exclusively, 10 (38%) sampled multiple habitats separately, 8 (31%) sampled multiple habitats in a pooled/random design, and 3 (12%) did not mention what habitats were sampled.
Control Sites	12 (46%) used 0 or 1 control sites, 8 (31%) used 2 or 3 control sites, and 6 (23%) used from 7 to 13 control sites.
Reference Sites	17 (65%) did not include a reference site, 3 (12%) used 1 reference, 6 (23%) used 2–5 reference sites
Standard of success	9 (35%) used a reference site as the standard of success for benthic macroinvertebrates. 17 (65%) compared restored to unrestored conditions (pre-restoration or control site)-typically assuming increased diversity, richness or B-IBI score is an improvement.
Habitat/substrate assessment	6 (23%) did not assess habitat/substrate. 8 (31%) usual visual assessments. 12 (46%) used quantitative measures to assess habitat.
Construction influence assessed/discussed	9 (35%) measured construction harm (through multiple measurements per year or multiple years, emphasizing the first year after construction and including pre-construction data or control sites.
Reference site parameters presented?	Of the 9 studies using reference stream conditions as the measure of success, 5 studies (55%) did not present any watershed attributes of the reference sites (watershed area, stream gradient, width, depth, discharge).
Land use assessed or discussed	50% discussed the land use of the catchment while 50% did not.

Miller et al. [18] conducted a quantitative meta-analysis on 24 studies (searching the literature using keywords including: restoration, rehabilitation, stream, river, invertebrates, macroinvertebrates, habitat, heterogeneity, channel reconfiguration), differentiating replicated from unreplicated studies, and concluded that heterogeneity enhancement projects had increased the macroinvertebrate richness but not diversity. Palmer et al. [17] compiled 18 such studies (12 of which overlapped the studies studied in Miller et al. [18]), and rather than combining the studies to test for statistical significance, evaluated each study independently. They reported no evidence of increased stream invertebrate diversity. Both Miller et al. [18] and Palmer et al. [17] cited the lack of robustness in the studies as a potentially important limitation. In particular, Miller et al. [18] noted the “(1) low quantity and poor quality of published biotic and abiotic data; (2) lack of rigorous study designs; (3) a dearth of replicated restoration efforts within physiographically similar areas”. The 26 studies reviewed by Miller et al. [18] and Palmer et al. [17] included a wide range of habitat enhancement actions and locations across the world, from Europe (14 studies), North America (10), Australia (1), and Japan (1). These studies represent genuine efforts to quantify the biological effect of the physical interventions to increase habitat complexity. However, the shortcomings of these studies were not previously systematically summarized.

2.1. Habitat Metrics

Although all studies sought to evaluate the biotic response to habitat heterogeneity enhancement projects, 12 studies (46%) quantitatively measured habitat alteration, with depth variability being the most common metric (Table 1). Other habitat heterogeneity metrics such as wood and organic matter retention, log or pool spacing, and bank erosion were used less frequently (Table 1).

Six studies (23%) did not report any data about habitats (in effect assuming that the restoration projects had actually increased habitat heterogeneity). If the hypothesis being tested is that enhanced habitat heterogeneity through restoration caused increased macroinvertebrate diversity, then it is important to quantitatively measure habitat heterogeneity to confirm whether or not the habitat heterogeneity has, in fact, been increased. This issue is further complicated by the absence of clear, agreed-upon, definitions of how to define and measure habitat heterogeneity. The importance of quantitatively and precisely assessing habitat heterogeneity was demonstrated by Laub et al. [49] who found that many unrestored (but non-channelized) urban streams had relatively high heterogeneity (measured through several specific metrics including variability in width, depth, velocity, thalweg profile, and bed sediment sorting) when compared to reference, forested sites, and that “restored” sites were often not more complex than unrestored sites. Thus, there is no a priori reason to assume that “heterogeneity enhanced” sites have more heterogeneity than unrestored urban stream sites.

Eight studies (31%) used visual estimates of grain size or adopted standard monitoring protocols for habitat quality (which include visual substrate estimates), such as the EPA Rapid Bioassessment Protocol (RBP) [50], the Ohio EPA Qualitative Habitat Evaluation Index (QHEI) [51], or the Bank Erosion Hazard Index (BEHI) [52]. These habitat quality metrics were developed primarily for quick implementation at a large number of sites, and they may not be well-suited for evaluation of individual restoration projects. They may not measure at the detail needed to determine causal relationships between habitat and macroinvertebrate communities. For example, visual substrate estimates have been shown not to be repeatable, and can lead to erroneous estimates. Visual estimates of substrate used in habitat modeling studies for the Physical Habitat Simulation Model (PHABSIM) overestimated median grain size when compared to the results from the scientifically established and repeatable method of pebble counts [53–55]. Whitacre et al. [56] compared six rapid assessment protocols and found statistically significant differences in results for nine out of ten basic habitat attributes such as sinuosity, percent pools, and median grain size. While rapid assessments may be useful for systematic monitoring of large spatial extents, there is little basis for their use in evaluating individual, reach-scale projects.

More comprehensively, Lisle et al. [57] questioned the basic premise of using rapid assessments to evaluate impairment of gravel-bed streams based on any protocol because any single metric is unlikely to reveal causative relations and channel condition can result from multiple pathways. Instead, Lisle et al. argued that channel condition be interpreted through the context of predictive mapping, site history and human influence. Standard metrics cannot be a substitute for hypotheses linking causes of impairment, project objectives, and restoration actions. For example, using bank erosion as a general measure of habitat quality (as is done in the EPA's Rapid Bioassessment Protocol and other rapid assessments [50]) either assumes that excess erosion is occurring at the site or that all bank erosion is undesirable. While channel incision and bank erosion are well-documented problems [58], particularly in urban settings, it may be misguided to seek more channel stability universally. To equate "failure" with bank erosion or the displacement of an in-channel structure wrongly assumes that successful stream restoration should create fixed, "stable" streams [59] even though ecological theory suggests, and experience on many rivers shows, that bank erosion can have many benefits, such as delivering spawning gravel to the channel [60], facilitating riparian vegetation succession [61,62], and providing habitat for early successional plants or disturbance-dependent species like bank swallows and yellow-billed cuckoo [63,64]. These examples call into question the underlying premise of "restoration" projects that seek to establish persisting forms rather than restoring the dynamic processes that would, in turn create such forms naturally.

2.2. Biological Metrics

Macroinvertebrate abundance, diversity, and composition (and indices of biological integrity derived from them) are some of the most commonly used measures of restoration success [65–70]. Of the 26 studies we reviewed, taxonomic richness of macroinvertebrates was the most common biological metric used to test the effects of habitat heterogeneity enhancements (21 studies, 81%). Abundance and density of macroinvertebrates were used in 17 studies (65%). Other diversity measures such as Shannon index or evenness indexes were used by 13 studies (50%). Composition of macroinvertebrate community or assemblage was used in nine studies (35%). Functional measures, such as functional feeding groups or trait composition were also used in nine studies (35%). Biological indices (such as the Benthic Index of Biotic Integrity B-IBI) were used in six studies (23%). Although all studies used taxonomic diversity, richness, or abundance of invertebrates as biologic indices, none presented explicit arguments why those metrics were relevant measures of success, but presumably these metrics were used as indicators of overall ecosystem health.

Of the 26 studies, five used the abundance of EPT (Ephemeroptera, Plecoptera, Trichoptera) taxa or the Benthic Index of Biotic Integrity (B-IBI) [69] as measures of success. The B-IBI was developed as a measure of water quality, based on the sensitivity of EPT taxa to water pollution. However, reach-scale habitat heterogeneity projects would not be expected to improve water quality per se unless, (1) excessive, locally-derived sediment was a limiting factor for sensitive species. Alternately, (2) habitat heterogeneity projects could increase the abundance of pollution-sensitive species by expanding or enhancing their desired habitats. In cases where log or rock structures were designed to create pools, macroinvertebrate communities could be expected to shift towards more pool-dwelling taxa. In cases where riffles were constructed, communities may shift towards more riffle-dwelling taxa. Without clear hypotheses developed before monitoring, monitoring pollution-sensitive taxa would yield essentially uninterpretable results. An increase in pollution-sensitive taxa could indicate either an improvement in water quality (1) or a change in morphological units that benefit the sensitive taxa (2). Since the abundance of sensitive taxa are being used as an indicator of water quality and are not typically themselves the restoration goal, knowing how and why changes occurred may be more important than simply documenting the change.

2.3. Controlling for Time, Space, and Variability

Of the 26 studies we reviewed, 16 (62%) included one year of post-project monitoring (Table 2). Sixteen studies (62%) had no pre-project monitoring and one study (4%) included more than one year of pre-project monitoring. Although the projects were designed specifically to create habitat heterogeneity (e.g., pools and riffles, velocity and depth heterogeneity, etc.), eight studies (31%) either sampled only riffles or did not specify what habitats were sampled. Ten studies (38%) sampled and analyzed biota from different habitats separately, and eight (31%) used either pooled or random designs that integrated samples from all habitat types. Nine studies (35%) used reference ecosystems as the standard of success. The other studies compared the restored reach to pre-restoration conditions or to a control site where legacy impacts persisted. Prolonged monitoring is necessary to discern the influence of restoration activities on biota [2,3,8]. For example, the disturbance caused by the restoration activity itself (i.e., vegetation clearing, dewatering the channel, compaction from heavy machinery in the channel, and grading the bed and banks) may decrease or increase biotic metrics such as abundance and diversity for days or decades, depending on severity of impact and rates of recovery. Many studies suggest the need to monitor at least several years in order to allow benthos to recover and to recolonize a restored reach [18]. The disturbance caused by restoration actions may increase diversity over intermediate time scales [71], so both pre-project monitoring and reference targets would be needed to interpret whether biotic differences following restoration are due to restored conditions, or to temporary effects.

Several decades of research have established how watershed conditions and position in the river system influence channel processes and forms [72] and ecosystem characteristics [73,74]. However, local conditions such as channel geometry can vary greatly over small distances or short time periods [75] because of changes in lithology and vegetation, tributary confluences, beaver activity, climate, or land use [76–78]. Therefore, it is essential to quantify the range of natural variability by compiling sufficient information from historical reference periods, multiple regional reference sites (using space for time substitution), or patterns, gradients, and processes that define reference conditions. Without understanding the natural variability of a system, meaningful restoration targets will be challenging to identify and monitoring data will be hard (or impossible) to interpret.

2.4. Relationship between Study Design and Biological Improvement

Seventy-eight percent of the studies that monitored for more than one year found increased diversity or richness. By contrast, 44% of the studies that sampled for only one year found increases. Similarly, studies that evaluated habitat directly were more likely to find increases (66%) than those using visual estimates or not presenting any habitat measures (50% in both cases) possibly reflecting a relationship between study effort and the significance of results, or perhaps indicating that more thorough evaluations followed better-planned restoration efforts. Again, studies using multihabitat sampling for macroinvertebrates either by pooled samples or differentiated habitats were also most likely to find improvement (75% and 60% respectively). By contrast, 20% of studies that sampled only one habitat (riffles) found increased diversity/richness.

3. Discussion

3.1. Are Reach-Scale Diversity and Abundance Universal Indicators of Success?

At the regional or global scale, changes in biodiversity may reflect overall ecological condition. However, in the context of reach-scale restoration, the “more is better” diversity assumption may not be appropriate. The use of macroinvertebrate diversity as an indicator is based on research linking species diversity to habitat heterogeneity, e.g., [79–81]. Increased habitat heterogeneity is assumed to provide more ecological niches for members of a community [82], provide refugia that stabilize predator-prey and host-pathogen dynamics, and generally support greater diversity and a more resilient ecosystem [17,83]. However, the universal applicability of diversity as a meaningful indicator

has not been established. Consider a few cases as thought experiments: an intermittent desert stream, a glacial outwash stream, and a headwater stream in old-growth forest. Human impacts could result in increased macroinvertebrate diversity (likely at the expense of native or endemic species) in all these cases.

In the intermittent desert stream example, perennialization resulting from urban development or irrigation return flow may allow new species to colonize desert washes. Richness and diversity have been strongly correlated with stream permanence in desert streams, though Feminella [84] found 7% of species were found only in intermittent streams, and presumably would be lost through perennialization. In the glacial outwash stream example, a warming climate may allow new species to establish, increasing diversity but threatening endemic specialists [85]. In the headwater stream example, logging of the old-growth forest could allow light penetration to the stream and increase diversity of habitats and species [86]. In all of these examples, diversity may be increased through additional generalist species, but communities of endemic specialists may be negatively affected. While all three thought experiments might be considered extreme cases, we know of no threshold for distinguishing “regular” streams (where more diversity might be an appropriate target) from these “extreme” cases, and thus find no basis for applying diversity as a universal indicator of success. Specifically addressing this issue in the context of restoration, Lepori et al. [35] found less macroinvertebrate species richness in Swedish reference streams than in either channelized or restored streams.

Food web interactions may be far more significant than diversity or abundance for influencing populations of the top predators for whose benefit restoration projects are undertaken. For example, in the Eel River, California, Power et al. [87] found that scouring winter floods promoted trophic interactions that produced more prey for steelhead, whereas drought years (or flow regulation) favored grazing macroinvertebrates that steelhead cannot eat. In this case, we can see that it is not necessarily the abundance or diversity of species that is important to top predators, but rather which species and energy pathways become dominant. In other cases, the removal of a key species can affect the whole ecosystem, like the Amazonian fish *Prochilodus mariae* whose removal from the Rio Las Marias in Venezuela altered nutrient cycling, sediment structure, and diatom and macroinvertebrate assemblages [88,89]. These examples suggest that restoration monitoring (and restoration objective themselves) require knowledge of the biophysical interactions that underpin the structure and dynamics of the target ecosystem.

Although restoration projects that aim restore habitat heterogeneity are supported by literature indicating that richness and diversity of biota increase with increasing diversity of habitats, changing the habitat away from its natural form and function is not typically an acknowledged goal. Leps et al. [70] conducted an exemplary and extensive assessment of 44 habitat heterogeneity projects in Germany. They quantitatively assessed 10 metrics of physical habitat and 33 biological metrics including measures of abundance, diversity, EPT taxa, and functional feeding groups. Leps et al. found that the restoration projects increased habitat heterogeneity, but did not produce detectable biological responses. What does that tell us? The study by Leps et al. included no reference sites, and ultimately had no method for defining target conditions at the 44 sites from across Germany. Were all 44 sites degraded with respect to abundance, diversity, EPT taxa, and certain functional feeding groups? Considering that the 44 sites were from different regions and watersheds of different sizes, then we would certainly expect target conditions to vary in significant ways. Indeed, an abundance of foundational ecological literature suggests that macroinvertebrate communities vary across the landscape and along river profiles [73,90]. We do believe that macroinvertebrates are useful indicators in many cases, but believe that the use of such indicators would be greatly improved by a priori establishment of target conditions for each site. The study by Leps et al. represents a tremendous effort in data collection and was conducted with robust methods and best practices. Ultimately, however, its results do not permit drawing conclusions about the success of habitat heterogeneity projects, because of what is probably considerable (but unmeasured) natural variability in conditions among sites. What we

learn from Leps and the papers of Miller and Palmer is that restoration evaluation may not be possible without a conceptual model that describes ecosystem functions and helps define target conditions.

3.2. Reference States and Best Practices

Appropriate regional reference sites may be challenging or impossible to find [91,92], especially for larger watersheds. Watershed position, regional climate, disturbance history, tectonic uplift rates, local geology, and many other factors will all influence stream processes and forms. However, physical controls such as channel slope, sinuosity, bed-material size, and precipitation are increasingly identifiable through automated monitoring and remote sensing [93]. That makes identification of appropriate reference sites more practical using GIS, and we expect analytical models will produce increasingly reliable predictions of watershed condition based these physical controls. Watershed comparisons have traditionally used a “paired basin” approach, in which “treated” are compared to “untreated” reference basins. Paired watershed experiments suffer from at least three problems (as reviewed by Reid et al. [94]) similar to those outlined above for restoration evaluations. First, “treatment” variables are usually only qualitatively characterized (e.g., “Managed vs. unmanaged” or “logged vs. unlogged”). Second, “control” treatments are never pristine. Third, even if untreated and treated watersheds have been matched with respect to aspect, area, slope, forest type, drainage density, and geological parent material, they may differ in subtle but important respects (e.g., structural orientation of bedrock, undetected ancient landslides whose scars are presently filled, and disease or fire history of vegetation). Comparisons of watershed outputs, such as total sediment yield at their mouths or changes in salmon escapement back to watersheds over the experimental period, are too noisy to reveal causality, particularly when observation records are short (less than decades). Several long-term and intensive monitoring efforts including the US Forest Service Experimental Watersheds, and the Intensively Monitored Watersheds Program, demonstrate the effort required to meaningfully quantify ecosystem change. The Caspar Creek Experimental Watershed has investigated logging practices and sediment yields for more than 50 years, finding that the sediment load doubled in response to logging, returned to initial levels 11 years after harvest, and then increased again a decade later as road crossings deteriorated during large storms [95]. Without a long-term record, assessing restoration and management actions would be confusing or misleading. Other long-term records and intensively studied watersheds provide important context on temporal and spatial variability.

Human alterations are widespread, and landscapes may still be responding to impacts from decades or centuries ago. Several reference streams are almost certainly required to adequately account for variability between streams and to test assumptions about variability in target conditions. Using an inappropriate reference stream (with different slope, drainage area, hydrologic regime, watershed position, etc.) can be misleading, and having only one reference stream is unlikely to help understand the natural range of processes and communities. Despite the challenges to use of reference sites, they provide a better basis for restoration targets than generic standards, which may not be locally appropriate. Of the 26 studies we reviewed, nine used reference streams, and of those, four presented information about the reference watershed. Of those four, two accepted considerable differences between fundamental watershed characteristics for reference and restored streams: in one study, the drainage area of the reference site was five times larger than the restored stream.

4. Conclusions

In the two decades since published calls for evaluation of stream restoration projects, most restoration projects remain un-evaluated. However, enough projects have now been evaluated that we can learn something from this experience. Metrics such as area or length of channel restored will probably not go away, but they are clearly inadequate to evaluate the success of restoration. Moreover, as reflected in recent experience, such metrics can greatly distort the types of projects undertaken for mitigation, as entrepreneurs seek to maximize mitigation credits. Many projects intended to enhance ecological complexity by increasing heterogeneity of habitats have been evaluated using abundance

and diversity of macroinvertebrate taxa. Interestingly, we found that the studies that conducted longer duration or more rigorous evaluations (e.g., direct habitat measurements, multi-habitat sampling for macroinvertebrates) were more likely to detect statistically significant increases in richness or diversity. This could be simply because the effects of the restoration required better monitoring to detect, but could also reflect a co-varying relationship between the resources available and the quality of both the restoration effort and the evaluation. Better-funded projects may be both better implemented and better evaluated, and thus more likely to produce a real effect, and more likely to detect the effect.

By contrast, some evaluations may be accomplished quite easily where the conceptual model of ecosystem dynamics is clearly established. Rood and colleagues [96] convincingly argued that dams and diversions in the Truckee River basin led to the collapse of the riparian ecosystem by altering the flow regime upon which cottonwood (*Populus fremontii*) and willow (*Salix exigua*) recruitment depended. Flows intended to restore the population of an endangered fish had the collateral benefit of recruiting riparian vegetation. In a straightforward evaluation, Rood et al. correlated seedling establishment and survival of seedlings with the changed flow regime. Finally, wetland- and riparian-dependent bird species were surveyed and compared to historical observations, demonstrating that several locally rare or extirpated species had returned with increased abundance [96]. The effort required for this evaluation was not extraordinary. What makes this an exemplary model is the clarity with which the conceptual model linking physical processes to habitat and species use was presented and evaluated.

We recommend a portfolio approach—combining knowledge from regional and historical reference streams, unrestored (control) sites, analytical models, and manipulative experiments—to define relevant target conditions. Prior to restoration, project designers and evaluators should develop conceptual models (which increasingly should include analytical reference states) of their ecosystems and consider success criteria carefully, in light of predictions generated from these models. Such an approach will add to the time and effort required for evaluation, but it is probably wiser to do a good job evaluating fewer projects than a poor job attempting many superficial evaluations.

Acknowledgments: We thank Mary Power and Kyle McKay for providing relevant ideas and references, special issue editor John Schwartz for insightful suggestions, and anonymous reviewers, whose suggestions improved the paper significantly. Research and manuscript preparation were supported by the Beatrix Farrand Fund of the Department of Landscape Architecture and Environmental Planning, and by a postdoctoral fellowship from the Ecuadorian National Secretariat of Higher Education (SENESCYT, contract 20120122) awarded to Blanca Rios-Touma.

Author Contributions: Zan Rubin, Blanca Rios-Touma, and Matt Kondolf conceived the study, analyzed the studies reviewed, reviewed relevant literature, and wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest. The funding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

BEHI	Bank Erosion Hazard Index
B-IBI	Benthic Index of Biological Integrity
EPA	US Environmental Protection Agency
EPA-RBP	US Environmental Protection Agency Rapid Bioassessment Protocol
EPT	Ephemeroptera, Plecoptera, and Trichoptera taxa
PHABSIM	Physical Habitat Simulation Model
QHEI	Ohio Qualitative Habitat Evaluation Index

References

1. National Research Council (U.S.), Committee on Restoration of Aquatic Ecosystems—Science, Technology, and Public Policy, Commission on Geosciences, Environment, and Resources. *Restoration of Aquatic Ecosystems: Science, Technology, and Public Policy*; Haworth Press: Philadelphia, PA, USA, 1992.
2. Kondolf, G.M.; Micheli, E.R. Evaluating stream restoration projects. *Environ. Manag.* **1995**, *19*, 1–15. [[CrossRef](#)]
3. Kondolf, G.M. Five elements for effective evaluation of stream restoration. *Restor. Ecol.* **1995**, *3*, 133–136. [[CrossRef](#)]
4. Kondolf, G.M. Learning from stream restoration projects. In Proceedings of the Fifth Biennial Watershed Management Conference, Ashland, OR, USA, 16–18 November 1994; Volume 3, pp. 133–136.
5. Downs, P.W.; Kondolf, G.M. Post-project appraisals in adaptive management of river channel restoration. *Environ. Manag.* **2002**, *29*, 477–496. [[CrossRef](#)]
6. Woolsey, S.; Capelli, F.; Gonser, T.O.M.; Hoehn, E.; Hostmann, M.; Junker, B.; Paetzold, A.; Roulier, C.; Schweizer, S.; Tieg, S.D.; et al. A strategy to assess river restoration success. *Freshw. Biol.* **2007**, *52*, 752–769. [[CrossRef](#)]
7. Palmer, M.A.; Bernhardt, E.S.; Allan, J.D.; Lake, P.S.; Alexander, G.; Brooks, S.; Carr, J.; Clayton, S.; Dahm, C.N.; Follstad Shah, J.; et al. Standards for ecologically successful river restoration: Ecological success in river restoration. *J. Appl. Ecol.* **2005**, *42*, 208–217. [[CrossRef](#)]
8. Rutherford, I.D.; Ladson, A.R.; Stewardson, M.J. Evaluating stream rehabilitation projects: Reasons not to, and approaches if you have to. *Aust. J. Water Resour.* **2004**, *8*, 57–68.
9. US Army Corps of Engineers. *Corps of Engineers Civil Works Direct Program Budget Development Guidance Fiscal Year 2016*; US Army Corps of Engineers: Norfolk, VA, USA, 2014.
10. Stakhiv, E. *Improving Environmental Benefits Analysis in Ecosystem Restoration Planning*; USACE, Institute for Water Resources: Alexandria, VA, USA, 2003.
11. Environmental Advisory Board to the Chief of the US Army Corps of Engineers. *Environmental Benefits and Performance Measures: Defining National Ecosystem Restoration and How to Measure Its Achievement*; Environmental Advisory Board to the Chief of the US Army Corps of Engineers: Norfolk, VA, USA, 2006.
12. McKay, S.K.; Linkov, I.; Fischenich, J.C.; Miller, S.J.; Valverde, L.J. *Ecosystem Restoration Objectives and Metrics*; U.S. Army Engineer Research and Development Center (ERDC), Environmental Laboratory: Athens, GA, USA, 2012.
13. Convertino, M.; Baker, K.M.; Vogel, J.T.; Lu, C.; Suedel, B.; Linkov, I. Multi-criteria decision analysis to select metrics for design and monitoring of sustainable ecosystem restorations. *Ecol. Indic.* **2013**, *26*, 76–86. [[CrossRef](#)]
14. Stewart, G.B.; Bayliss, H.R.; Showler, D.A.; Sutherland, W.J.; Pullin, A.S. Effectiveness of engineered in-stream structure mitigation measures to increase salmonid abundance: A systematic review. *Ecol. Appl.* **2009**, *19*, 931–941. [[CrossRef](#)] [[PubMed](#)]
15. Krebs, C.J. The experimental paradigm and long-term population studies. *Ibis* **1991**, *133*, 3–8. [[CrossRef](#)]
16. Bernhardt, E.S.; Palmer, M.; Allan, J.D.; Alexander, G.; Barnas, K.; Brooks, S.; Carr, J.; Clayton, S.; Dahm, C.; Follstad-Shah, J.; et al. Synthesizing U.S. river restoration efforts. *Science (Washington)* **2005**, *308*, 636–637. [[CrossRef](#)] [[PubMed](#)]
17. Palmer, M.A.; Menninger, H.L.; Bernhardt, E. River restoration, habitat heterogeneity and biodiversity: A failure of theory or practice? *Freshw. Biol.* **2010**, *55*, 205–222. [[CrossRef](#)]
18. Miller, S.W.; Budy, P.; Schmidt, J.C. Quantifying Macroinvertebrate Responses to In-Stream Habitat Restoration: Applications of Meta-Analysis to River Restoration. *Restor. Ecol.* **2010**, *18*, 8–19. [[CrossRef](#)]
19. Wohl, E.; Angermeier, P.L.; Bledsoe, B.; Kondolf, G.M.; MacDonnell, L.; Merritt, D.M.; Palmer, M.A.; Poff, N.L.; Tarboton, D. River restoration. *Water Resour. Res.* **2005**, *41*, W10301. [[CrossRef](#)]
20. Reynoldson, T.B.; Norris, R.H.; Resh, V.H.; Day, K.E.; Rosenberg, D.M. The reference condition: A comparison of multimetric and multivariate approaches to assess water-quality impairment using benthic macroinvertebrates. *J. N. Am. Benthol. Soc.* **1997**, *8*, 833–852. [[CrossRef](#)]
21. Biggs, J.; Corfield, A.; Grøn, P.; Hansen, H.O.; Walker, D.; Whitfield, M.; Williams, P. Restoration of the rivers Brede, Cole and Skerne: A joint Danish and British EU-LIFE demonstration project, V—Short-term impacts on the conservation value of aquatic macroinvertebrate and macrophyte assemblages. *Aquat. Conserv. Mar. Freshw. Ecosyst.* **1998**, *8*, 241–255. [[CrossRef](#)]

22. Ebrahimnezhad, M.; Harper, D.M. The biological effectiveness of artificial riffles in river rehabilitation. *Aquat. Conserv. Mar. Freshw. Ecosyst.* **1997**, *7*, 187–197. [[CrossRef](#)]
23. Edwards, C.J.; Griswold, B.L.; Tubb, R.A.; Weber, E.C.; Woods, L.C. Mitigating effects of artificial riffles and pools on the fauna of a channelized warmwater stream. *N. Am. J. Fish. Manag.* **1984**, *4*, 194–203. [[CrossRef](#)]
24. Friberg, N.; Kronvang, B.; Ole Hansen, H.; Svendsen, L.M. Long-term, habitat-specific response of a macroinvertebrate community to river restoration. *Aquat. Conserv. Mar. Freshw. Ecosyst.* **1998**, *8*, 87–99. [[CrossRef](#)]
25. Gerhard, M.; Reich, M. Restoration of Streams with Large Wood: Effects of Accumulated and Built-in Wood on Channel Morphology, Habitat Diversity and Aquatic Fauna. *Int. Rev. Hydrobiol.* **2000**, *85*, 123–137. [[CrossRef](#)]
26. Gørtz, P. Effects of stream restoration on the macroinvertebrate community in the River Esrom, Denmark. *Aquat. Conserv. Mar. Freshw. Ecosyst.* **1998**, *8*, 115–130. [[CrossRef](#)]
27. Harrison, S.S.C.; Pretty, J.L.; Shepherd, D.; Hildrew, A.G.; Smith, C.; Hey, R.D. The effect of instream rehabilitation structures on macroinvertebrates in lowland rivers. *J. Appl. Ecol.* **2004**, *41*, 1140–1154. [[CrossRef](#)]
28. Jähnig, S.C.; Lorenz, A.W. Substrate-specific macroinvertebrate diversity patterns following stream restoration. *Aquat. Sci.* **2008**, *70*, 292–303. [[CrossRef](#)]
29. Jähnig, S.C.; Lorenz, A.W.; Hering, D. Restoration effort, habitat mosaics, and macroinvertebrates—Does channel form determine community composition? *Aquat. Conserv. Mar. Freshw. Ecosyst.* **2009**, *19*, 157–169. [[CrossRef](#)]
30. Jähnig, S.C.; Lorenz, A.; Hering, D. Hydromorphological parameters indicating differences between single- and multiple-channel mountain rivers in Germany, in relation to their modification and recovery. *Aquat. Conserv. Mar. Freshw. Ecosyst.* **2008**, *18*, 1200–1216. [[CrossRef](#)]
31. Jungwirth, M.; Moog, O.; Muhar, S. Effects of river bed restructuring on fish and benthos of a fifth order stream, Melk, Austria. *Regul. Rivers Res. Manag.* **1993**, *8*, 195–204. [[CrossRef](#)]
32. Laasonen, P.; Muotka, T.; Kivijärvi, I. Recovery of macroinvertebrate communities from stream habitat restoration. *Aquat. Conserv. Mar. Freshw. Ecosyst.* **1998**, *8*, 101–113. [[CrossRef](#)]
33. Larson, M.G.; Booth, D.B.; Morley, S.A. Effectiveness of large woody debris in stream rehabilitation projects in urban basins. *Ecol. Eng.* **2001**, *18*, 211–226. [[CrossRef](#)]
34. Lemly, A.D.; Hilderbrand, R.H. Influence of large woody debris on stream insect communities and benthic detritus. *Hydrobiologia* **2000**, *421*, 179–185. [[CrossRef](#)]
35. Lepori, F.; Palm, D.; Brännäs, E.; Malmqvist, B. Does restoration of structural heterogeneity in streams enhance fish and macroinvertebrate diversity? *Ecol. Appl.* **2005**, *15*, 2060–2071. [[CrossRef](#)]
36. Lester, R.E.; Wright, W.; Jones-Lennon, M. Does adding wood to agricultural streams enhance biodiversity? An experimental approach. *Mar. Freshw. Res.* **2007**, *58*, 687–698. [[CrossRef](#)]
37. Moerke, A.H.; Gerard, K.J.; Latimore, J.A.; Hellenthal, R.A.; Lamberti, G.A. Restoration of an Indiana, USA, stream: Bridging the gap between basic and applied lotic ecology. *J. N. Am. Benthol. Soc.* **2004**, *23*, 647–660. [[CrossRef](#)]
38. Muotka, T.; Laasonen, P. Ecosystem recovery in restored headwater streams: The role of enhanced leaf retention. *J. Appl. Ecol.* **2002**, *39*, 145–156. [[CrossRef](#)]
39. Muotka, T.; Paavola, R.; Haapala, A.; Novikmec, M.; Laasonen, P. Long-term recovery of stream habitat structure and benthic invertebrate communities from in-stream restoration. *Biol. Conserv.* **2002**, *105*, 243–253. [[CrossRef](#)]
40. Nakano, D.; Nakamura, F. Responses of macroinvertebrate communities to river restoration in a channelized segment of the Shibetsu River, Northern Japan. *River Res. Appl.* **2006**, *22*, 681–689. [[CrossRef](#)]
41. Negishi, J.N.; Richardson, J.S. Responses of organic matter and macroinvertebrates to placements of boulder clusters in a small stream of southwestern British Columbia, Canada. *Can. J. Fish. Aquat. Sci.* **2003**, *60*, 247–258. [[CrossRef](#)]
42. Pedersen, M.L.; Friberg, N.; Skriver, J.; Baattrup-Pedersen, A.; Larsen, S.E. Restoration of Skjern River and its valley—Short-term effects on river habitats, macrophytes and macroinvertebrates. *Ecol. Eng.* **2007**, *30*, 145–156. [[CrossRef](#)]
43. Purcell, A.H.; Friedrich, C.; Resh, V.H. An assessment of a small urban stream restoration project in northern California. *Restor. Ecol.* **2002**, *10*, 685–694. [[CrossRef](#)]

44. Roni, P.; Bennett, T.; Morley, S.; Pess, G.R.; Hanson, K.; Slyke, D.V.; Olmstead, P. Rehabilitation of bedrock stream channels: The effects of boulder weir placement on aquatic habitat and biota. *River Res. Appl.* **2006**, *22*, 967–980. [[CrossRef](#)]
45. Rosi-Marshall, E.J.; Moerke, A.H.; Lamberti, G.A. Ecological responses to trout habitat rehabilitation in a Northern Michigan stream. *Environ. Manag.* **2006**, *38*, 99–107. [[CrossRef](#)] [[PubMed](#)]
46. Sarriquet, P.-E.; Bordenave, P.; Marmonier, P. Effects of bottom sediment restoration on interstitial habitat characteristics and benthic macroinvertebrate assemblages in a headwater stream. *River Res. Appl.* **2007**, *23*, 815–828. [[CrossRef](#)]
47. Tullos, D.D.; Penrose, D.L.; Jennings, G.D.; Cope, W.G. Analysis of functional traits in reconfigured channels: Implications for the bioassessment and disturbance of river restoration. *J. N. Am. Benthol. Soc.* **2009**, *28*, 80–92. [[CrossRef](#)]
48. Walther, D.A.; Whiles, M.R. Macroinvertebrate responses to constructed riffles in the Cache River, Illinois, USA. *Environ. Manag.* **2008**, *41*, 516–527. [[CrossRef](#)] [[PubMed](#)]
49. Laub, B.G.; Baker, D.W.; Bledsoe, B.P.; Palmer, M.A. Range of variability of channel complexity in urban, restored and forested reference streams. *Freshw. Biol.* **2012**, *57*, 1076–1095. [[CrossRef](#)]
50. Barbour, M.T.; Gerritsen, J.; Snyder, B.D.; Stribling, J.B. *Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers*; United States Environmental Protection Agency, Office of Water: Washington, DC, USA, 1999.
51. Ohio Environmental Protection Agency, Division of Surface Water. *Methods for Assessing Habitat in Flowing Waters: Using the Qualitative Habitat Evaluation Index (QHEI)*; Ohio Environmental Protection Agency: Columbus, OH, USA, 2006.
52. Rosgen, D.L. A practical method of computing streambank erosion rate. In Proceedings of the Seventh Federal Interagency Sedimentation Conference, Reno, NV, USA, 25–29 March 2001; Volume 2, pp. 9–15.
53. Wolman, M.G. A method of sampling coarse river-bed material. *EOS Trans. Am. Geophys. Union* **1954**, *35*, 951–956. [[CrossRef](#)]
54. Kondolf, G.M. Application of the pebble count: Notes on purpose, method, and variants. *J. Am. Water Resour. Assoc.* **1997**, *33*, 79–87. [[CrossRef](#)]
55. Kondolf, G.M.; Li, S. The pebble count technique for quantifying surface bed material size in instream flow studies. *Rivers* **1992**, *3*, 80–87.
56. Whitacre, H.W.; Roper, B.B.; Kershner, J.L. *A Comparison of Protocols and Observer Precision for Measuring Physical Stream Attributes 1*; Wiley Online Library: Hoboken, NJ, USA, 2007.
57. Lisle, T.E.; Buffington, J.M.; Wilcock, P.R.; Bunte, K. Can rapid assessment protocols be used to judge sediment impairment in gravel-bed streams? A commentary. *JAWRA J. Am. Water Resour. Assoc.* **2015**, *51*, 373–387. [[CrossRef](#)]
58. Shields, F.D.; Lizotte, R.E.; Knight, S.S.; Cooper, C.M.; Wilcox, D. The stream channel incision syndrome and water quality. *Ecol. Eng.* **2010**, *36*, 78–90. [[CrossRef](#)]
59. Bennett, S.J.; Simon, A.; Castro, J.M.; Atkinson, J.F.; Bronner, C.E.; Blersch, S.S.; Rabideau, A.J. The evolving science of stream restoration. In *Stream Restoration in Dynamic Fluvial Systems*; Wiley Online Library: Hoboken, NJ, USA, 2013; pp. 1–8.
60. Buer, K.; Eaves, J.N. *Middle Sacramento River Spawning Gravel Study*; Department of Water Resources: Sacramento, CA, USA, 1984.
61. Piégay, H.; Cuaz, M.; Javelle, E.; Mandier, P. Bank erosion management based on geomorphological, ecological and economic criteria on the Galaure River, France. *Regul. Rivers Res. Manag.* **1997**, *13*, 433–448. [[CrossRef](#)]
62. Florsheim, J.L.; Mount, J.F.; Chin, A. Bank erosion as a desirable attribute of rivers. *Bioscience* **2008**, *58*, 519–529. [[CrossRef](#)]
63. Gardali, T.; Holmes, A.L.; Small, S.L.; Nur, N.; Geupel, G.R.; Golet, G.H. Abundance patterns of landbirds in restored and remnant riparian forests on the Sacramento River, California, USA. *Restor. Ecol.* **2006**, *14*, 391–403. [[CrossRef](#)]
64. Golet, G.H.; Brown, D.L.; Carlson, M.; Gardali, T.; Henderson, A.; Holl, K.D.; Howell, C.A.; Holyoak, M.; Hunt, J.W.; Kondolf, G.M.; et al. Successes, failures and suggested future directions for ecosystem restoration of the middle Sacramento River, California. *San Franc. Estuary Watershed Sci.* **2013**, *11*. [[CrossRef](#)]
65. Brudvig, L.A. The restoration of biodiversity: Where has research been and where does it need to go? *Am. J. Bot.* **2011**, *98*, 549–558. [[CrossRef](#)] [[PubMed](#)]

66. Ruiz-Jaen, M.C.; Mitchell Aide, T. Restoration success: How is it being measured? *Restor. Ecol.* **2005**, *13*, 569–577. [[CrossRef](#)]
67. Bonada, N.; Prat, N.; Resh, V.H.; Statzner, B. Developments in aquatic insect biomonitoring: A comparative analysis of recent approaches. *Annu. Rev. Entomol.* **2006**, *51*, 495–523. [[CrossRef](#)] [[PubMed](#)]
68. Resh, V.H. Which group is best? Attributes of different biological assemblages used in freshwater biomonitoring programs. *Environ. Monit. Assess.* **2008**, *138*, 131–138. [[CrossRef](#)] [[PubMed](#)]
69. Kerans, B.L.; Karr, J.R. A benthic index of biotic integrity (B-IBI) for rivers of the Tennessee Valley. *Ecol. Appl.* **1994**, *4*, 768–785. [[CrossRef](#)]
70. Leps, M.; Sundermann, A.; Tonkin, J.D.; Lorenz, A.W.; Haase, P. Time is no healer: Increasing restoration age does not lead to improved benthic invertebrate communities in restored river reaches. *Sci. Total Environ.* **2016**, 557–558, 722–732. [[CrossRef](#)] [[PubMed](#)]
71. Connell, J.H. Diversity in tropical rain forests and coral reefs. *Science* **1978**, *199*, 1302–1310. [[CrossRef](#)] [[PubMed](#)]
72. Petts, G.E.; Amoros, C. *The Fluvial Hydrosystem*; Springer: Dordrecht, The Netherlands, 1996.
73. Vannote, R.L.; Minshall, G.W.; Cummins, K.W.; Sedell, J.R.; Cushing, C.E. The river continuum concept. *Can. J. Fish. Aquat. Sci.* **1980**, *37*, 130–137. [[CrossRef](#)]
74. Ward, J.V.; Stanford, J.A. The serial discontinuity concept of lotic ecosystems. *Dyn. Lotic Ecosyst.* **1983**, *10*, 29–42.
75. Waters, M.R.; Haynes, C.V. Late Quaternary arroyo formation and climate change in the American Southwest. *Geology* **2001**, *29*, 399–402. [[CrossRef](#)]
76. Montgomery, D.R. *Process Domains and the River Continuum*; Wiley Online Library: Hoboken, NJ, USA, 1999.
77. Polvi, L.E.; Wohl, E.E.; Merritt, D.M. Geomorphic and process domain controls on riparian zones in the Colorado Front Range. *Geomorphology* **2011**, *125*, 504–516. [[CrossRef](#)]
78. Wohl, E. *Mountain Rivers Revisited*; John Wiley & Sons: Hoboken, NJ, USA, 2013; Volume 19.
79. Simpson, E.H. Measurement of diversity. *Nature* **1949**, *163*, 688. [[CrossRef](#)]
80. MacArthur, R.H.; Wilson, E.O. *Theory of Island Biogeography. (MPB-1)*; Princeton University Press: Princeton, NJ, USA, 2015; Volume 1.
81. Ricklefs, R.E.; Schluter, D. Species diversity: Regional and historical influences. In *Species Diversity in Ecological Communities*; University of Chicago Press: Chicago, IL, USA, 1993; pp. 350–363.
82. Warfe, D.M.; Barmuta, L.A.; Wotherspoon, S. Quantifying habitat structure: Surface convolution and living space for species in complex environments. *Oikos* **2008**, *117*, 1764–1773. [[CrossRef](#)]
83. Bell, S.; McCoy, E.D.; Mushinsky, H.R. *Habitat Structure: The Physical Arrangement of Objects in Space*; Springer Science & Business Media: Berlin, Germany, 2012; Volume 8.
84. Feminella, J.W. Comparison of benthic macroinvertebrate assemblages in small streams along a gradient of flow permanence. *J. N. Am. Benthol. Soc.* **1996**, *15*, 651–669. [[CrossRef](#)]
85. Jacobsen, D.; Milner, A.M.; Brown, L.E.; Dangles, O. Biodiversity under threat in glacier-fed river systems. *Nat. Clim. Chang.* **2012**, *2*, 361–364. [[CrossRef](#)]
86. Hernandez, O.; Merritt, R.W.; Wipfli, M.S. Benthic invertebrate community structure is influenced by forest succession after clearcut logging in southeastern Alaska. *Hydrobiologia* **2005**, *533*, 45–59. [[CrossRef](#)]
87. Power, M.E.; Parker, M.S.; Wootton, J.T. Disturbance and food chain length in rivers. In *Food Webs*; Springer: New York, NY, USA, 1996; pp. 286–297.
88. Flecker, A.S. Ecosystem engineering by a dominant detritivore in a diverse tropical stream. *Ecology* **1996**, *77*, 1845–1854. [[CrossRef](#)]
89. Taylor, B.W.; Flecker, A.S.; Hall, R.O. Loss of a harvested fish species disrupts carbon flow in a diverse tropical river. *Science* **2006**, *313*, 833–836. [[CrossRef](#)] [[PubMed](#)]
90. Rice, S.P.; Greenwood, M.T.; Joyce, C.B. Tributaries, sediment sources, and the longitudinal organisation of macroinvertebrate fauna along river systems. *Can. J. Fish. Aquat. Sci.* **2001**, *58*, 824–840. [[CrossRef](#)]
91. White, P.S.; Walker, J.L. Approximating nature's variation: Selecting and using reference information in restoration ecology. *Restor. Ecol.* **1997**, *5*, 338–349. [[CrossRef](#)]
92. Whittier, T.R.; Stoddard, J.L.; Larsen, D.P.; Herlihy, A.T. Selecting reference sites for stream biological assessments: Best professional judgment or objective criteria. *J. N. Am. Benthol. Soc.* **2007**, *26*, 349–360. [[CrossRef](#)]

93. Gilvear, D.; Bryant, R. Analysis of Remotely sensed data for fluvial geomorphology and river science. In *Tools in Fluvial Geomorphology*, 2nd ed.; John Wiley & Sons: Chichester, UK, 2016; pp. 103–132.
94. Reid, L.M.; Dunne, T.; Cederholm, C.J. Application of sediment budget studies to the evaluation of logging road impact. *J. Hydrol.* **1981**, *20*, 49–62.
95. Keppeler, E.T. Sediment production in a coastal watershed: Legacy, land use, recovery, and rehabilitation. In *Proceedings of the Coast Redwood Forests in a Changing California: A Symposium for Scientists and Managers*, Santa Cruz, CA, USA, 21–23 June 2011; Gen. Tech. Rep. PSW-GTR-238; Pacific Southwest Research Station, Forest Service, U.S. Department of Agriculture: Fresno, CA, USA, 2012, 2012; pp. 69–77.
96. Rood, S.B.; Gourley, C.R.; Ammon, E.M.; Heki, L.G.; Klotz, J.R.; Morrison, M.L.; Mosley, D.; Scopettone, G.G.; Swanson, S.; Wagner, P.L. Flows for floodplain forests: A successful riparian restoration. *BioScience* **2003**, *53*, 647–656. [[CrossRef](#)]



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).