Decision
by Karl Cottenie, 2019-11-10 00:00
Manuscript: https://www.biorxiv.org/content/10.1101/611939v1.abstract

This preprint merits a revision

I have now read the two reviews of the preprint "On the efficacy of restoration in stream networks: comment and critique" by Murray-Stoker. Based on their comments, I have again read the preprint, and here is my recommendation. I think, though, that my recommendation might not be very satisfying for all parties involved. I hope, though, that I provide enough justification to convince the reader that a new approach would be beneficial.

Both reviewers point out the dual objectives of this preprint: on the one hand it provides a reanalysis of existing data from a previous study based on a different opinion on study design and analysis decisions, and on the other hand a critique of the motivation or intent of these differences. I agree with both reviewers that re-analysis of existing data in papers show the strength of the current practice of open and reproducible research. However, I also agree with one of the reviewers that the manuscript lacks some structure to easily convey the differences of the three papers in question. And their recommendation of a table with three columns, one for the original paper, one for the erratum, and one for this preprint would be useful. Each row would be a comparison item, which could be what studies were included, what design decisions were made, what statistical decisions were made, what results and conclusions were made for each of these 3 approaches. That will make it easier for the reader to compare and contrast, and then also decide where they stand on each of these decisions.

I also agree with one of the reviewers, though, that attributing negative intent to the original authors seems harsh and even unnecessary. They shared their data to make all this re-analysis even possible, so I think it is very unlikely that they intentionally made decisions that would bias their results and thus their conclusions. I think that what we are seeing here is that open science exposes more of the messiness of scientific practice with subjective decisions that have always been present, but now are out in the open and can be questioned and potentially corrected. This productive debate can lead to better or different analyses and conclusions, and it is up to the reader to decide which one is the most convincing one.

This is important for me, because after reading all the different opinions, I think that there is an inherent flaw with the current analysis method. If I understand it correctly, for most of the restored streams, there is a comparison between the restored stream and an adjacent stream. (I tried to find the original data mentioned in the preprint, but for some reason the link did not work.) If this is in essence a comparison between an experimental and control condition, and the study wants to combine results from different experiments, I think that the appropriate statistical approach should be a meta-analysis. This has several advantages. For each experiment, you will be able to directly compute an effect size (which was also advocated by one of the reviewers) for each study (in essence treatment - control, potentially scaled by sample size and/or standard deviation). This direct comparison will ensure that the reach dependency is included in the analysis but not as a random effect, and it will also reflect in the figures the true power of the statistical design. Right now when you look at the figures, this paired nature of the design is not at all represented. Secondly the meta-analytical approach would avoid one of the points raised by the author: what studies to include and which ones not to include. If you work with effect sizes in a meta-analysis, one of the assumptions being made by the different authors (whether bank stabilization and in-channel manipulations treatments are strong enough) will actually not be an assumption but a test. In the meta-analysis, you could test whether the stronger impacts result in larger effect sizes, or whether the different types of treatments (reforestation etc) have different effect sizes. Meta-analysis techniques were developed to exactly address these types of questions. Finally, by computing the effect size, you will avoid having to include the random stream identity, and since this is a difference, not only is the interpretation more directly, but also more likely to result in a statistical model with easier model assumptions.

I realize that this recommendation for an additional or different analysis is one of the reasons that started this debate and series of manuscripts. However, I hope that my arguments outlined above will convince somebody that the meta-analytical approach is superior to the analyses from the previous articles and preprint, and that this might provide a more detailed and useful analysis and associated conclusions.

Response: All of the above criticisms are fair and have been addressed whenever possible in the revised manuscript. One issue brought up by all reviewers is the tone, which has been extensively revised.

My understanding of the effect size calculations was to basically compare the effects in each of the models; I did this using partial eta^2 values, as this was a metric that could easily be used across studies (i.e. it only required degrees of freedom and F statistics and I could not replicate results from the initial study). Based on your comments and those of Reviewer 1, this was my understanding of the "meta-analysis" technique to use. This allows for comparisons of the models across studies and goes beyond just simple comparisons of statistical significance but actual measures of the effect sizes. By adding this to the broader analytical approach, readers can see differences in statistical meaning and interpretation (e.g. use of Type I vs. Type III sums of squares, statistical significance of an effect or interaction) and incorporate the ecological relevance of any effect. This was a great suggestion, as it shows how the models could be over- or under-estimating different effects and could identify low-power/high-effect size factors to focus on for future study.

Regarding the paired structure of the design in the figures, I replicated the figure presentation of Swan and Brown (2017, 2018). The paired design is at the reach level, so headwaters have paired restored-unrestored reaches and mainstems have restored-unrestored reaches. What was included as a random effect was the stream identity, but each restored-unrestored reach was nested within an individual stream. I have added some more information regarding the sampling design to help clarify the misunderstanding.

Reviewer 1
Reviewed by Eric Harvey, 2019-10-21 22:23
General comments
The author offers here a criticism and re-analysis of the data from a previous study published by Swan and Brown. The author findings seem to be strikingly different from the one in the original study casting doubts on the quality of the data processing and analyses performed in the original study. In the original study, the authors had found an overall positive effects of habitat restoration on diversity and temporal stability in headwater but not mainstem sites in accordance with their expectations. Here the author show, after correcting for some issues with the analyses, that there are in fact no detectable effects of habitat restoration on any of the diversity metrics regardless of whether the site is located up or downstream. The author then conclude on the potential causes of this absence of effect.

The exercise proposed here by the author is on its own definitely valuable. Reproducing results from published studies is an important way to validate results and improve reproducibility of scientific studies. At best, it should initiate discussions or debates and eventually lead to a more informed consensus on the state of our understanding and knowledge on a specific issue (here the efficiency of habitat restoration in river systems). It should also help us to identify issues with the way we sometime use and interpret statistics so that everyone end up being better off from the process (i.e., a constructive process). Here, however, I would argue that some of the technical issues found should clearly have been found during the reviewing process, but that is another question. Sadly, despite all this, as I will expose in more details below, the article does give an impression of a public *vendetta*, making important accusations against specific authors. At the very least, PCI should invite Brown and Shawn to write a reply.

As is probably obvious from what I wrote so far, I enjoyed reading this manuscript. I also think that it is very clearly written. However I have a few main questions and concerns before it can published and some minor concerns below that I hope the author, editor and PCI will find useful:

1) The first criticism exposed by the author in the Introduction is more philosophical and open to debate I think than the other ones which are more technical. The author argues that in comparing restored vs. unrestored sites, Swan and Brown should have taken into account that not all restoration treatments should be expected to lead to the same outcome or to influence headwater versus mainstem sites the same way. I must admit that I am on the fence here. While I agree that not all habitat restoration treatments are the same, I still see values in making a broad hypothesis such as the one made by Swan and Brown: All else being equals, headwaters will be more affected by habitat modifications than mainstems because of fundamentals X and Y. To me what the author is getting at here is the idea of context-dependency. It ultimately would suggest that Swan and Brown were simply lucky to find results matching their main hypotheses because of the specific treatments they were looking at (with their specific effects not cancelling each other out) and that re-sampling the general pool of the different restoration types could lead to different results (i.e., "identity effect" matters). In that sense the author brings here an important note of caution about making sweeping generalizations or inference from a limited dataset. Basically here the author is building his own set of expectations for a different, but still interesting study (or more like a follow-up). If the goal was to make a more general point, then

why a focus on one study and not just a more general concept and synthesis (or Forum) sort of paper?
Response:

2) The tone: I am not arguing that the author should have just said nothing when he found those issues (of course, not!) - but why going all-in public and writing a whole article clearly targeting specific authors? As a reviewer I am missing background information here to assess the situation and so it's very hard for me to understand some of the accusations. The author goes as far as accusing them of *Questionnable Research Practices*, which is an important offense because it could suggest that they manipulated data or analysis to mislead readers on purpose. This accusation also comes with no evidences of this. I don't know all the details but did the author contact Swan and Brown first? The concluding message of the article upon re-analyzing the data is quite interesting, but is loss in what seems to me like a public *vendetta*. As mentioned above I think it is essential that Swan and Brown are given the opportunity to submit a reply.
Response: I have taken into consideration the tone and made revisions on that aspect of the manuscript. As this manuscript was written as a comment paper or letter to the editor, one would find that the tone is not too dissimilar from other comments or letters critiquing previous research. Those comments and letters are just as much a "public vendetta" as what has been written and submitted here, and comments and letters to the editor are commonly published in ecological and evolutionary journals.

To add more context, the impetus for this pre-print is because the Swan and Brown (2017) article was rapidly involved in a broad discussion within freshwater ecology groups, and I wanted to add caution to the original study. It was and still is important work within the freshwater ecology, metacommunity, and restoration ecology disciplines, but those discussions were based on results that were not supported upon closer scrutiny. To put it simply, I did not want a study with fundamental flaws to be discussed and used in the literature without necessary critique.

For more clarity, background, and context, I contacted the handling editor of the initial paper (Swan and Brown 2017) soon after reading the article once issued, and I also contacted Swan and Brown directly at that time, as suggested by the editor. I shared with them my R code and the reasoning behind the re-analysis at the time, and they had all of the R code for months prior to submitting an erratum, which, to state clearly, was only issued because I noticed errors in the initial study. I thought it was a very interesting study building on previous work done by Swan and Brown, and further building on links between metacommunity theory and restoration ecology, but there were still errors. I submitted a letter to the editor that was rejected, and correctly, as I did not give enough detail on the statistical side of the story. After rejection, I further raised my concern of irreproducible results and discrepancies between described and conducted analyses via email correspondence to the editor, but no further action was taken and my concerns were ignored. I was informed Swan and Brown were required to submit the R code with the erratum; if one analyzed their code with no changes to the analyses (e.g., no Type III sums of squares, no correction to address homogeneity of variance), you do get the results in the erratum. The issue is the results presented in the erratum retained statistical errors and, when analyzed properly, results are considerably different. What may appear as a "vendetta" is actually a search for the correct results; for over two years, incorrect results have been in the public domain, informing research (cited > 20 times) and policy (Hildrebrand et al. 2018, Chesapeake Bay Trust), and even upon releasing an erratum there are still issues. I respect Christopher Swan and Bryan Brown, as their research in metacommunity ecology has been very influential to the ecological community, but this particular project has required proper revision and correction for over two years. I do not consider pushing for the correct results to be in the public domain after repeated failures in the review process for over two years a "vendetta."

The tone in the pre-print was interpreted as over-the-top, so it has been revised. I would just add that the tone is not too dissimilar from that of many other comments and letters to the editors in journals like Ecological Applications, which I used to model and structure the original manuscript. I am not saying that because other comments had a specific tone my tone is acceptable, but rather providing the background and perspective for how the manuscript was structured. My stating of questionable research practices could require more exact evidence, but to change an analysis without noting the change in the text of a manuscript or a supplement is at best negligent. For at least some evidence of questionable research practices, in the initial study (Swan and Brown 2017), it was explicitly stated for the temporal variation analysis that using a Jaccard index and an incidence matrix yielded stronger results than a Bray-Curtis index and abundance matrix; upon evaluating the erratum (Swan and Brown 2018) and associated R code for the temporal variation analyses, it reads as if the Bray-Curtis index, abundance matrix, and group centroid were used (initial study and erratum text), but it was a Jaccard index, incidence matrix, and spatial median (associated R code in the supplement). This is some evidence, explicitly stated by the authors before the erratum was even conceived, that different dissimilarity indices and community matrix types would affect the strength of the analysis, and this was known to the authors. In my view, that borders on or crosses the line into questionable research practices. I would further argue that intent is not always relevant, as you can unknowingly conduct questionable research practices and deceive the readers of the work and achieve the same ends as researchers actively seeking to deceive or manipulate their data.

This manuscript was written as a letter to the editor or a comment paper, and not a standard research article; comment papers and letters to the editor appear in many ecology and evolution journals, and the aim of this critique is the same as any other letter to the editor. If those are not considered public vendettas, I do not think this could be considered a "public vendetta." I welcome Swan and Brown being invited to submit a reply, as they were made aware of my concerns over two years ago, and R code for my re-analysis has also been in the public domain for months. I have brought up these concerns for two years, and the scientific review process has failed on multiple occasions, so this is the last avenue through which concerns can be presented, discussed, and debated.

3) Several of the criticisms made by the author rely on a hard stance on statistics that is not always completely justified I think. That Brown and Swan should have used a Type III analysis is clear and sound, but many are highly skeptical of data-transformations and would even argue that linear models tend to be quite robust to violations of their assumptions (or then why not using a generalized linear model rather than a transformation?). I am not taking a side here as this is outside my field of expertise. But if a simple log transformation changes the results so much, could it simply be that fundamentally the main treatment effects are statistically significant but biologically meaningless? On this, it would be interesting for the author to present effect sizes for each set of models to give the reader an idea. This could open up on a more

general discussion, I think, in ecology about how we tend to interpret statistical outputs based only on a p-value. I feel like perhaps if the signal was that strong, the changes made by the author would not have changed the results that much. Perhaps overall this is more a constructive tale of caution about how we interpret the statistical significance vs. the biological importance of the processes we attempt to measure and quantify.
Response: I do not want to put myself forth as an expert in statistics (I am largely self-taught, as formal statistical education is inadequate in ecology and evolution), but there are some issues that were not adequately addressed by Swan and Brown (2017, 2018) but I feel are addressed in my manuscript. Although linear models are generally robust to violations of the assumptions, the issue comes with the highly-unbalanced design when all sites are considered. Unequal sample sizes could bias the F statistic, depending on where the greater variation lies. The effects of unequal sample sizes would be compounded further by using Type I sums of squares, rather than Type III sums of squares, particularly when interactions are present and, in fact, when interactions are expected. The differences in the results between my re-analysis and those of Swan and Brown (2017, 2018) are predominantly due to the use of Type III sums of squares, which are essentially required for the analysis to be valid. For example, the full and reduced models by Swan and Brown (2018) presented in Table 2 had identical model structure and parameters to Swan and Brown (2018), with the only difference being the use of Type III sums of squares. To end with a note on whether data should or need to be transformed, my philosophy is the assumptions of the test must be adequately met before any consideration and interpretation of the results; if the data violated the test assumptions, results derived from that test should not be used.

I agree that a presentation of effect sizes could generate a more productive discussion than just looking at the p-values. P-values only indicate whether something is statistically significant, not whether it is ecologically relevant or meaningful. The full issue and discussion of the utility of p-values goes beyond what could be addressed in this manuscript but is something that needs discussion among the greater ecological community of researchers. My aim with using p-values was just to illustrate the differences in interpretation based on statistical significance, which highlighted the discrepancies between papers; use of p-values was not and should not be used as evidence for strength of an effect.

I have added effect sizes into the revised manuscript, which I again want to agree was a great suggestion. I calculated the effect sizes as partial eta^2, which allows for calculations requiring only degrees of freedom and F statistics; this permits easier translation across manuscripts as the initial study (Swan and Brown 2017) used standard sums of squares, while part of the erratum (Swan and Brown 2018) and my re-analysis used a maximum likelihood approach. F statistics and degrees of freedom were the only values consistently reported, and, as the analyses from the initial study could not be replicated, this was the only method to maintain consistency.

Although documentation of the partial eta^2 method is detailed in the revised manuscript, it is taken from Cohen (1973), where:

$$\eta = \frac{df_{between} \times F}{df_{between} \times F + df_{within}}$$

Minor comments

Introduction

The Introduction is clear and well written. The author gives enough information so that a naive reader can get and idea of Swan and Brown hypotheses. The author also explains clearly his perceived issues with the way data were analyzed in Swan and Brown.

Methods

[91] does the author meant a PERMANOVA? Otherwise it's not clear how these multi-variate dissimilarity analyses were performed.

Response: The modified Gower distances were calculated for each restored-adjacent pair, and those distance values were then included as the response variable in a standard ANOVA (as per Swan and Brown 2017). I have revised this sentence to be clearer.

[88:95] Why a Gower for spatial dissimilarity and a Bray-Curtis for temporal dissimilarity? (I understand that the author cannot respond for Swan and Brown). It's also not clear to me why no transformations were imposed on the abundance matrix. It's been shown many times that some transformations (e.g., hellinger) can really improved detections of patterns in the data by weighting the disproportionate effects of rare and very abundant species.

Response: I only used the modified Gower index to replicated the analysis by Swan and Brown (2017). My understanding for using modified Gower is to consider n-fold changes in abundance as a difference in composition. Working in rivers and streams, you can have Chironomidae comprising the vast majority of the abundance of some communities, but other important taxa from a restoration and bioassessment standpoint, Ephemeroptera (mayflies), Plecoptera (stoneflies), and Trichoptera (caddisflies), are generally of lower abundance.

I cannot provide any further justification for choice of dissimilarity index or lack of transforming the abundance matrix, as I was just following the methods written by Swan and Brown (2017).

[113] would not ln-transforming the Gower base 5 dissimilarity values totally alter the interpretation of the dissimilarity index?

Response: To my knowledge and based on my understanding, ln-transforming the modified Gower dissimilarity values would not "totally alter" the interpretation. The exact values would be different, but the interpretation should be unchanged (e.g. the dissimilarity between sites A and B was higher than the dissimilarity between sites A and C).

[123] Here the "nlme" package is cited - unless I am missing something I think the author should then use the more specific "linear mixed effect models" rather than ANOVA.

Response: I fitted the models using the 'lme()' function, following Swan and Brown (2018). The only difference between doing an ANOVA on an lme() fitted model and a standard lm() or aov() in R is that the mixed models use maximum likelihood approach and not sums of squares. lme() is commonly used for specifying model structure and then conducting an ANOVA.

[160:168] The idea suggested at the end is quite interesting but rather speculative in light of what the author says about his capacity to test his hypothesis.

Response: One of the issues with the experimental design is that treatments were not assessed factorially, so isolating the individual contributions of any one treatment is not possible;

however, in the absence of this knowledge, reducing the variation in what treatments were applied so there is greater consistency among sites would better address how restoration affects the communities. Additionally, as this manuscript is written as a formal comment on previous manuscripts, my aim was to identify areas of concern to inform future studies.

Results

The difference in the results, even for the same 'full sites' data are quite stricking. Unless I missed something, it seems like for the 'full sites' data, the only differences between the original study by Swan and Brown and the re-analysis here is that 1) Response variables were transformed to meet model assumptions and 2) a Type III analysis was used. Am I correct? Those transformations were performed "to better meet the assumption" - did the transformation actually made the data fit the assumptions? Also how does a square-root transformation on diversity, an ln transformation on richness and Gower dissimilarity index actually affect interpretations of those response variables? Why not using a generalized-linear model instead? I am asking because those transformations might be one of the main reasons for that stricking difference in the results with the original study, but sometimes transformation are known to have undesirable impacts on variable distribution and interpretation. If a simple transformation changes the results so much, could it simply be that fundamentally the main treatment effects are statistically significant but biologically meaningless? On this, it would be interesting for the author to present effect sizes for each set of models.
Response: You are correct that the changes made in the full sites analysis was the transformation of variables and Type III sums of squares; I also removed a single, unpaired site from the re-analysis. Data transformations helped the data to better meet model assumptions (i.e. 'perfect' homogeneity of variance was still not achieved, but distributions of residuals were improved), but the same trends from the analyses would be observed without transformed data. Is the use of data transformation absolutely necessary? No, not in isolation; however, given the imbalance of sample sizes, I wanted to reduce the heterogeneity in variances and better meet model assumptions whenever possible. The difference in results between Swan and Brown (2017. 2018) and my re-analysis is predominantly due to the use of Type III vs. Type I sums of squares.

To aid with the interpretation of transformed variables for statistical analyses, untransformed values were presented in the figures. I have had mixed advice of presenting untransformed values when variables were transformed for analysis, but my position is that untransformed variables are easier to visualize and interpret; transforming variables is only done to meet, or better meet, test assumptions.

Again, I do agree with the presentation of effect sizes. I had not considered gathering those data, but it is a good suggestion. Effect size calculations have been incorporated.

Reviewer 2
Mariana Perez Rocha

This is a very well written manuscript addressing issues found in Swan and Brown 2017, 2018. The author points out in this manuscript the main flaws found in Swan and Brown 2017 (and later the Erratum 2018): experimental design, the improper use of statistical analyses, and the discrepancies between what was written in the methods and what was actually conducted. That being said, the author provides secondary sound analysis of the data utilized in Swan and Brown 2017, and based on the new sets of results, the author delivers more accurate interpretations of the results in the context of metacommunity ecology framework and stream restoration. I believe the following comments would improve the quality and flow of this manuscript.

-Title: the title would read and sound better if added "…….. : comments and critiques".
Response: Thank you for the suggestion, the title has been revised.

-Abstract: This section is the most important part of a manuscript and yet it appears to be very short and lacking crucial information. I'm not sure what the target journal to be submitted is, but I would suggested expanding this section a little. In the end of the Abstract I was missing one of the main points stated throughout the text: the author's conclusions after re-analyzing Swan and Brown 2017 data and the ecological implications of the "Swan and Brown 2017 misleading analysis and interpretations".
Response: Thank you for this comment. I modelled the length and structure from the abstracts of comment papers and letters to the editor in Ecological Applications and The American Naturalist. I wanted to send this comment to Ecological Applications, which does not have abstracts for comments/letter to the editor, but that avenue is no longer a viable option. Given I am not restricted to the formatting of Ecological Applications, I do have more space in which to include the main points, and, as you have highlighted, and I have provided the main conclusions and implications in the abstract of the revised manuscript.

-Results: It would be very helpful for the reader including a table or a figure summarizing/comparing the main findings in Swan and Brown 2017, 2018 and the findings after the re-analysis of the data.
Response: Thank you for this suggestion. I have incorporated a table the presents the results from the initial study (Swan and Brown 2017), the erratum (Swan and Brown 2018), and if results were consistent upon the full or revised sites re-analysis.

-Lines 151-153: These lines correspond to hypothesis stated by the author. But this hypothesis itself do not appear before in the text. I strongly recommend to the author clearly including/stating such hypothesis (i.e. effectiveness of local restoration) in the last paragraph of the Introduction section where the author is bringing up "why" the data found in Swan and Brown 2017 should be re-analyzed; or making clear if the author here are just re-analyzing Swan and Brown 2017 hypotheses (as I found later in the lines 242-243 "I then evaluated the same hypothesis proposed by Swan and Brown….").

Response: Thank you for this comment, it was really helpful. I certainly needed to have greater clarity in the manuscript, and the revised version should help in this regard. I set and evaluated my own hypothesis, that stream-channel manipulations would have a more consistent effect relative to the effect of riparian reforestation between headwaters and mainstems, with stronger effects of restoration in headwaters relative to mainstems; however, I did not clearly state this in the introduction, and it was first explicitly stated in the discussion. I have now stated my hypothesis in the last paragraph of the introduction, adding that the hypothesis was used to guide site selection in the re-analysis. Given I was re-analyzing data from a previous study, I could not effectively isolate the effects of in-stream and riparian reforestation treatments. Limitations of my hypothesis notwithstanding, without examining Table 2 in Swan and Brown (2017) and wanting to test how restoration treatments could confound the results, I would not have conducted the re-analysis and discovered discrepancies between published results and what is evident upon re-analysis. I do hope the hypothesis serves to inform future stream restoration experiments, either by myself or other researchers.

An area of confusion is that my hypothesis retains similarity to that of Swan and Brown (2017, 2018), specifically that effects of restoration will be dependent upon stream location (i.e. headwaters or mainstems); this hypothesis has both theoretical and empirical support (Heino et al. 2003, Journal of Animal Ecology 72:425–434; Altermatt 2013, Aquatic Ecology 47:365–377; Heino 2013, Oecologia 171:971–980), including previous research by Swan and Brown (Brown and Swan 2010, Journal of Animal Ecology 79, 571–580). Swan and Brown (2017, 2018) had strong support for their hypothesis, and I agreed with the majority of the study; I only disagreed with the incorporation of all sites regardless of the combination of applied restoration treatments.

-Lines 249-251: I recommend to the author excluding these lines. Indeed, Swan and Brown 2017, 2018 results were quite misleading, and what has been presented in the 2018 erratum does not quite match the code provided. The author has clearly addressed these issues throughout the manuscript text. However, stating that "Swan and Brown 2017, 2018 demonstrated questionable research practices" is quite harsh to be said in a manuscript and might sound too offensive.
Response: Thank you for this suggestion, it has been revised. You are correct that it sounds too offensive and accusatory, particularly as any motives and intentions are currently not known. I have revised the manuscript to a more amicable tone.

-Lines 240-258: Over these lines one can find the concluding remarks. However, I miss in the end of this section a more 'concluding remark'. This section was quite repetitive on what the author has previously stated throughout the manuscript. I believe the author show include in this section a better suit of 'ecological implications and prospective suggestions' related to issues presented in this manuscript.
-Response: This is a fantastic suggestion. You are correct that it is repetitive, and a section titled "Ecological Implications & Prospective Suggestions" would be more appropriate, and this has been added in the revised manuscript. In the revised manuscript, I focus on the ecological interpretation and consequences of the results, with suggestions for future research in river and stream restoration projects.