PLOS ONE

Dear Editor Bolboacă and Reviewers,

Thank you for the opportunity to revise and resubmit our manuscript. We appreciate the time and effort taken by the editor and reviewers and think their comments and feedback have significantly improved our manuscript. Below, we have outlined our response to the editor and reviewer comments in table format, with editor/reviewer comments in the left column and our corresponding response on the right column. Note that page #s refer to the manuscript without track changes. To summarize, we have clarified the argument throughout, significantly re-organized the entire manuscript, and filled in gaps in discussing the results and discussion.

We thank you for your time and consideration and look forward to your response.

Sincerely,

Authors

|  |  |  |
| --- | --- | --- |
| **#** | **Itemized Review Comments** | **Response** |
|  | *Editor Comments* |  |
| 1 | Differences between fields are highlighted since its introduction (Hirsch, J.E.: An index to quantify an individual’s scientific research output. Proceedings of the National Academy of Sciences, USA 102(46), 16569- 16572 (2005)). The scientific literature is reach in this domain and it is unclear why a new article is needed or which this article bring in. | The main contribution of the article is the analysis of social factors, specifically gender and sole authorship, contributing to within and between disciplinary differences. While Hirsch's original article describes how between disciplinary differences may occur, prior work does not evaluate the balance of how sole authorship and gender relate to these differences. The random effects within-between models also allow us to model both between and within discipline effects. The field effects finding is also new as it shows that the disciplinary effects vary by field with variation in both focal factors. With your thoughtful comments and the reviewers' numerous comments highlighting the manuscript's strengths in mind, we have edited the manuscript throughout to emphasize its main contributions. |
| 2 | It is unclear which type or article is this. | We have incorporated the appropriate subheadings as recommended, so that the resulting manuscript follows the structure of an article according to journal guidelines. We apologize for the confusion caused by the prior structure. |
| 3 | Play attentions to misspellings ("ypotheses") | Thank you for bringing this to our attention. We have carefully copyedited the document. |
| 4 | Do not start a sentence with an abbreviation. | We have removed the abbreviation. |
| 5 | (Abstract) It is unclear when the study was conducted, which h-index, how the scientist were selected etc. The results presented in this section are not specific. Please also define “apples-to-oranges”. | We have provided more detail on the data coverage, the scientist selection, and defined apples-to-oranges. We also added to the discussion of the results in the abstract. |
|  | (Introduction) |  |
| 6 | Differences in demographic characteristics, like race and 4 gender, capture processes of racism and sexism" differences in which way? | We have clarified by revising to the following on page 1: Inequalities experienced by minoritized and women scholars point to processes of social closure or exclusionary practices that limit the opportunities for some scholars and not others across disciplines. |
| 7 | Appropriately present the state of the art. To say "A large body of research describes the factors that contribute to inequality in academia [see 9, 10, for reviews]." is uninformative. The same comment for listing references and sending the reader to read more. | We have deleted this sentence as the next sentence summarizes the current literature with more specificity. |
| 8 | "differences exist in the distribution of resources to disciplines" please be specific. | We have revised for clarity on page 2: differences exist in the distribution of resources to disciplines, such as the prioritization of the natural sciences and engineering in federal funding and the unequal distribution of resources, like salaries, to departments, the proxies for disciplines on university campuses |
| 9 | "To understand factors that contribute to differences in scholar’s h-index, or Hirsch index, we use data on over 50,000 high-performing scientists across 174 disciplines in the United States" this information belongs to the methods section. | We have removed this sentence. |
| 10 | Line 43-60 belongs to Methods, Results and Conclusions. | We have removed this section. |
| 11 | "less is known 41 about how within and between discipline differences affect h-index scores " disagree. | We have clarified the contribution in the following revision on page 2: While prior work examines the \*h\*-index by field, discipline, or gender, less is known about how gender and sole authorship contribute to \*both\* within \*and\* between discipline differences in \*h\*-index scores. |
| 12 | End this section with the aim of your study. | We have added more detail regarding the aims of the study in the first section of the Introduction. |
| 13 | "Inequality in science" must be integrated in the Introduction | We have integrated "Inequality in Science" into the Introduction. (see also #2) |
| 14 | "publications produced by a scientist translates into tangible resources, like salary raises" this is not universally true. | We have clarified this point by revising this sentence as follows on page 3: How many and the type of publications produced by a scientist often translates into tangible resources, like salary raises and the job security of tenure, and less tangible resources, like prestige. |
| 15 | "publication remains a site of persistent inequality" in which ways? | We have both softened this claim and made it more specific: “publication remains a potential site of stratification in terms of both the number of publications scholars accrue over their careers and the quality of those publications.” |
| 16 | "While women have made significant gains in many academic fields, " duplicated information. | This duplication has been removed. |
| 17 | "The risks of quantification " and "The case of the h-index" should be part of Introduction. | We have integrated these sections into the Introduction. (see also #2 and #13) |
| 18 | Some hypotheses were already demonstrated (ex. H3a). | Our initial thought was to include this as a baseline but agree in hindsight it is unnecessary as a hypothesis. We have deleted it. |
|  | (Materials and methods) |  |
| 19 | the h-index is that in Scopus? Unclear. | The h-index was constructed by the authors of the dataset using Scopus data. We cite the dataset on page 7. |
| 20 | The composition of the five fields is unclear. | The composition of the fields is provided in the supplementary table S1. We have changed the sorting of that table so that disciplines are sorted by fields and then alphabetically to assist readers in understanding this composition. |
| 21 | It is unclear how career length was evaluated. | We clarify career length on page 7. |
| 22 | Define STEM abbreviation. | We have defined the STEM abbreviation. |
| 23 | lines 319-329 is duplicated information (see Table 1). | We have removed most of this paragraph and only included a single sentence reference to Table 1. |
| 24 | It is unclear how were identified the disciplines with a higher share of women or those with higher share of sole authorship. | We have substantially rewritten the description of the REWB models to clarify how we are measuring variables at both the individual and disciplinary level on pg. 9. In this case, we simply take the mean of the individual variables across disciplines to determine these shares. |
| 25 | It seems that you evaluated sex not gender. | We have thought seriously about this issue and believe that it is a complicated one. Thank you for bringing it to our attention. The data used to estimate gender is based on sex, but name effects, as we understand the current sociological literature, is a gendered process. Current literature on this topic refers to name estimation as "gender" or "gendered" and the packages estimating gender, such as gender in R used here, discuss their estimation techniques as inferring gender (Blevins and Mullin, 2015; Dworkin et al., 2020). |
|  | (Results) |  |
| 26 | "We begin by analyzing a partition of the variance in the h-index within and between 366 disciplines in a null multilevel model. The percentage of the total variation in the 367 h-index that occurs between disciplines is given by the intraclass correlation coefficient 368 (ICC) of this model." this information belongs to methods. | After considering the comment here, we feel that the selected paragraph properly belongs in the results. The partitioning of the variance is not simply a technical issue but is related directly to our hypotheses, specifically H3a. Therefore, we have chosen to keep the paragraph as part of the results. |
| 27 | Table 2 - % should stay wither in definition of row or column NOT in the body of the table. | We have removed percentage marks from the body of the table and placed them in the column headers to indicate that percentages sum within columns. |
| 28 | Do not discuss in this section the results. | We have moved part of the results that were more of a discusssion and interpretation to a new discussion section as discussed below. |
| 29 | Discussion section is missing. | Our conclusion contained much of the information consistent with a discussion section. We have added a discussion section and moved this information to the new discussion section. |
| 30 | Which is the generalizability of the reported results? | A discussion of generalizability is now included in the discussion section (see page 11-12). |
| 31 | Which is the usefulness of the reported results? | A discussion of the usefulness of the results is now included in the discussion section. |
| 32 | Which are the limitations of your study? | Data limitations are now included in the discussion section alongside the generalizability discussion (see page 12). |
|  | (Conclusion) |  |
| 33 | "The h-index, or Hirsch Index, is a widely used metric used for performance evaluation or 428 quality valuation in the sciences and across the academy. This research aimed to 429 contribute to the literature on bibliometrics and inequality in science by examining both 430 within and between discipline differences in the h-index. We used REWB models to 431 predict the h-index for high-performing scholars in 174 disciplines. " this is a summary and duplicate information. Please delete. | We have deleted this duplication. We have carefully reviewed the manuscript to eliminate any other redundancies that are a result of disicplinary-specfic stylistic conventions to better correspond with the PLOS style. |
|  | *Reviewer #1* |  |
| 34 | Thank you for the opportunity to review the manuscript “Inequality in measuring scholarly success: Variation in the h-index within and between disciplines”. | Thank you for reviewing this manuscript. Your comments are greatly appreciated. |
| 35 | Among the positive features of this work, I want to note a proper review of the relevant scientific literature on this problem. The authors logically structured and filled the subsections with the manuscript: Inequality in science, Inequality between disciplines and fields, The risks of quantification, The case of the h-index. At the same time, the limitations of quantitative analysis are well and consistently described, especially regarding the comparison of different disciplines. | Thank you for these positive comments about the structure of our approach and for our discussion of its strengths and weaknesses. |
| 36 | The researchers employ a substantial amount of data to carry out their investigation (data from over 50,000 high performing scientists across 174 disciplines). The use of variation in h-score indices across scholars using random effects within-between (REWB) models ensures methodological rigour and validity of this study. The authors' conclusions are clear, logical, properly presented and supported by data. | Thank you for providing such a thorough, but concise summary of our manuscript. |
| 37 | I found it particularly interesting to learn that about a third of the h-index variation is due to cross-disciplinary rather than intra-disciplinary differences. And also that more frequent sole authorship is associated with a lower h-index. | Thank you for identifying some specific points of interest. These have informed our revision. |
| 38 | I have no advice or comments regarding the scientific content of this manuscript, I noticed only two typographical errors: 257 hypotheses 381 misperceive | Thank you for identifying these typographical errors. We have fixed them and carefully reviewed the manuscript for any additional errors. |
|  | *Reviewer #2* |  |
| 39 | This study explored the inequality in using the h-index to measure scholarly success. The research design is rigorous and the findings are thought-provoking. The manuscript is in general qualified for publication, except for the several issues below to address. | Thank you for reviewing this manuscript and for your positive assessment of it. Your comments are greatly appreciated. |
| 40 | 1. It seems weird to pose research questions at the very beginning (the first paragraph) of the introduction section, even without a quick and brief review of past studies. | Thanks for this point. We have placed the research questions after the second paragraph that provides an overview of relevant literature. This sets up the research question in a much better way. |
| 41 | 2. The structure of the main body of the literature review is not clear. Although headings are used, the several factors contributing to inequality are intertwined, which may make readers confused. Are the factors of gender, discipline, collaboration, etc. paralleled or nested? I think the authors should re-organize this section. | We have addressed this issue by more closely following the PLoS structure. We hope that the main factors are appropriately parallel. |
| 42 | 3. The last paragraph of p11 is not clear enough. To be specific, what is the relationship between feminization and sole-authorship, as is revealed by Model 3 and 4? The interactive effect is very interesting, but the authors failed to explain it in a detailed and convincing way. | We have edited this paragraph to clarify that we are not analyzing an interactive effect but rather observing how a model coefficient changes with the addition of other variables. In this case, the initial negative association of the feminization variable with h-index scores was a spurious byproduct of the association between feminization of a discipline and the tendency for sole-authorship. We have clarified this issue in the revision (see page 10). |
| 43 | 4. The discussion of the results should be expanded. The current discussion is too simple, though the results are interesting. | We have now added a separate discussion section to the paper at the request of the editor. We have included in this section a summary of our results that links them back to each of the hypotheses in the front end of the paper. |
| 44 | 5. There are several typos (e.g. p7 line 258 “yphtheses”). Proof-reading should be done before submission of the revised edition. | Thank you for this comment. We have proof-read the manuscript closely and also fixed some LaTex compiling issues. |

*References*

Blevins, C., & Mullen, L. (2015). Jane, John... Leslie? A Historical Method for Algorithmic Gender Prediction. DHQ: Digital Humanities Quarterly, 9(3).

Dworkin, J. D., Linn, K. A., Teich, E. G., Zurn, P., Shinohara, R. T., & Bassett, D. S. (2020). The extent and drivers of gender imbalance in neuroscience reference lists. Nature neuroscience, 23(8), 918-926.