Revision round #2: Revision needed

Recommender: François Munoz

The reviewers and I are pleased with your revision. One of the reviewers suggested a few changes. I would be happy if you could take them into account in a new version. This is a minor revision. I think we can then complete the process quickly.

reply#1: Thank you and the two reviewers Matthias Grenié and François Massol for appreciating our work and effort we put in the revision. We have addressed and incorporated the last comments – please, see below our replies. Changes implemented in the main text are highlighted in blue.

Review by Matthias Grenié

I would like to take the oppportunity of this review to thank the authors for their revised work, as well as their detailed answers to my previous comments. I think their manuscript is now stronger both methodologically and conceptually given the additional discussions, and I don't have further comments to make. I think this manuscript is great contribution to the debate on confirmatory versus novel research. I'd also like to thank the recommender for giving the opportunity to review this manuscript.

reply#2: Thanks for providing very useful feedback in the previous round of revision that stimulated us to improve the quality and clarity of the scientific message contained in the manuscript.

Review by Francois Massol

The revision of the paper by Ottaviani et al. is a clear improvement -- kudos to the authors for managing this so well! I have no major remark.

reply#3: Thanks a lot to you as well – you offered very thoughtful comments during the revision of our manuscript that helped us to refine the analytical framework and contextualization of the research.

Since I have spotted some items that might need some rewriting, here is the list.

General minor comment: once printed on paper, the figures are difficult to read. This might be a dpi issue -- my advice would be to make sure all figures are as readable as possible given that they contain quite a number of character strings (notably figs 1 and 3).

reply#4: Thanks for spotting this, but we would prefer to wait till the typesetting stage before implementing further changes – the journal may also require other variations in the figure formatting.

Line 48: Connor -> Conner

reply#5: Changed.

Lines 79-81: While the current wording is not wrong, I would probably argue that the relationship between IF and the frequency of terms has a lot to do with the link between prestige (or perceived prestige) of journals and the way the editors in those same journals encourage or discourage certain types of writing. The "journal", by itself, is not really creating this association -- in my opinion, this is all in the mind of the beholders, here the beholders being the editors and reviewers working for these journals.

reply#6: Very good suggestion – we have added a sentence in **L81-83**.

Line 98: The wording does not seem to imply that (scoring 1 for at least one novelty word) => (scoring 1 for novelty). Or, in other words, it is well explained how you count occurrence of words in abstracts, but not how you decide that the abstract is in the category "novel" and/or "confirmatory". My understanding is that it is the case if you have at least one mention of these words. It might be good to explain this in a more explicit fashion.

reply#7: Thanks – this was probably not clear. We made it more explicit (**L105**).

Pages 7 and 10: About the IF, I cannot find the information on the year of IF, and especially whether it was kept the same for all papers independently of their publication year or whether it was actualized -- for all the papers, did you use the IF as given by Clarivate/ISI at the time of publication or as the IF was in 2017 (or 2023, or any recent year actually)?

This question has a practical consequence: on page 10, you show a positive relationship between the age of the paper and its IF. I'd suggest two completely different interpretations based on how the IF was reported in your dataset:

* if the IF reported for a paper is the one of its journal at the time it was published, then it might be that all IFs have declined over time, so that only old papers can claim a high IF.

* if the IF reported for a paper is kept fixed for all papers belonging to the same journal (and taken at the most recent value), then it might mean that IFs have generally increased with

time but that journals publish more papers when they have lower IFs (and thus there would be an over-abundance of old papers in now high-IF journals).

reply#8: Thanks, this was indeed not clear. We used actualized Journal Impact Factor, and we now clarified it in **L130-131**. This way, we could track/observe the year-by-year link between the frequency of terms' use and Journal Impact Factor.

Page 7: Regarding the linear model on the IF -- does it mean you used a Gaussian distribution? ("linear model" now means so many different things, I'm a bit lost, sorry). If so, did you use the raw IF or a transformation (like log IF) in order to correct for the asymmetry/impossibility of becoming negative? (using a Gaussian distribution with a variable that will never be negative means you are wagering some bet [in the likelihood of the model] on the possibility that the variable takes negative values, even when you know it mechanically cannot do so).

reply#9: It is true that Gaussian distribution may not be the most appropriate given Journal Impact Factor cannot assume negative values. However, since the model validation for a linear model was OK, we preferred to keep it simple rather than opting for more complex distributions (such as Gamma). This approach is recommended in Zuur & Ieno (2016), which we generally followed as a roadmap in all regression analyses (see **L109**). We also added more explanation of it at **L132-136**.