

Dear Authors,

Thank you for submitting a revised version of this manuscript within which you have made extensive changes based on the reviewers comments and outlined in detail where and why you have/have not made the amendments. This paper is very much moving in the right direction.

There are some additional elements that require amendment and/or clarification before publication which I have outlined here that focus on both the statistics and some minor revisions such as references missing, clarifications or related to minor spelling, punctuation and grammar errors. I would be grateful if you could clarify and/or amend the following - the line numbers refer to the 'tracked changes' version of the manuscript.

Minor revisions

Ln 10 - should this say 'diversify' rather than 'diversity'?

Ln 15 – I believe this should say 'offered' not 'offers'

Ln 35 - is there a reference you can add to support the phrase 'leaky pipeline'?

Ln37 – Can you briefly outline the structural challenges faced within this sentence?

Ln 38 - I'm not sure if it is the formatting but I think the citation might be missing a parenthesis.

Ln46-80 – Is there are reason these data are so Germany/USA focused? It needs some explanation or alternatively, consideration of data from other countries. You explain in Line 567 why the data are focused on Germany and the USA but I feel it would be beneficial to explain earlier.

Ln 57 - Check capitalisation of PhD.

Ln79 – Refers to scientists. Is that accurate or should it be academics? This depends on which data you are referring to.

Ln93 – Is this just relevant to scientific conferences?

Ln95 – I believe this should say 'organise' rather than 'organising'

Ln100-101 – Needs a reference to support the claim that not participating is the only option. How do you know this?

Ln 119 - It may be helpful to outline the discipline this work is from to distinguish it from the geosciences.

Ln 136 – Year is missing from citation

Ln 138 - are all of these conferences based in the Global North, South or a mixture? Might be worth acknowledging any potential bias in the data.

Ln 147 – 151 – Write 'Table' out in full rather than the shortened 'Tab.'

Ln 158 – Remove 'we' and rephrase this sentence

Ln 193 - 'these groups' - please can you be specific and name them

Ln 205 – 'English' needs a capital letter

Ln 206 - I believe this should say "stated they..." rather than "stated to..."

Figure 2 caption – this should be Figure 2a and Figure 2b rather than panel A&B. Please check all other figures to ensure parity.

Ln 312 – What does 'not shown' mean?

Ln 321 - Is this perceived loneliness? I'm not convinced that the current quote supports that idea. Do you have another quote that does support this? If not, I don't think you can make this claim.

Ln 344 - Does it increase mental load? Is there evidence to support this? If not, this should be removed.

Ln 383 – Are the spelling errors of 'to/too' verbatim? If so, please indicate this i.e. (sic) after the quote. There is irregular punctuation on this line.

Ln 342 to Ln 436 - this whole section should be reconsidered as only towards the end does it become apparent that the suggestions you are making are based on evidence from the survey. It currently mostly reads as your own suggestions and feels quite out of place. I would suggest adding some qualitative quotes into the section as results and evidence of the claims. I recognise that there are two quotes within here, but it would be beneficial to get a sense of the comments that led you to these suggestions. Once you have rewritten this, the recommendations then may feel out of place so it may be worth keeping this section as a results section and moving the discussion/recommendations elsewhere or signposting to the checklist. Please review the structure of this section once you have rewritten it.

Ln666 – I believe 'wether' should be spelled 'whether'

Ln 667 – Where are these data from?

Ln670 – 'for several years' is a better phrase than 'since several years'

Ln 676 - Liang & Prior (2025) - first names are not necessary.

Please ensure a full spelling, punctuation and grammar check is carried out prior to resubmission.

Statistics

It is recommended that you talk to someone who is comfortable with statistics for this last bit of tightening up. For anything you say, it's testable to see if it's actually observable (i.e. not just by chance). There is not a need to emphasise the whole distribution statistics, which for them mainly now say - there's no difference, no result to see here, in most of the figures. There are things that are relatively clear in the data, as outlined below, however, it is suggested that you just need to be a bit more tactical / explicit in your use of the statistics to support your observations.

- The tests in Fig. 2 & Fig. 3 & Fig. 4 & 6 all show $p > 0.05$, in other words the distributions cannot be distinguished as different with the χ^2 test. So, it's not possible to interpret any difference between them. Suggest that this is added to the figure caption if not already in the text.

- Section 3.2 - interpretation/claims appear to be in line with the statistical tests
- Some things are perhaps surprising i.e. in Fig. 4 if considering took 'yes' and 'no' it looks like it should be statistically significant - probably including the two don't know options diluted this.
- Please ensure that you are clear about exactly what you are testing each time as chi-squared could be used for a number of purposes. It is suggested that comparing specific columns would also be useful e.g. in Figure 6 supporting a statement like '*more childcare was done by partners than family and friends*' [3x2 Chi-squared] or '*friends were used more often by scientists with 3+ friends as opposed to those with 0 or 1*' [a binomial test would work here, or 2x2 Chi-squared] and you are encouraged to consider whether χ^2 is the best approach to compare the whole distribution shape.
- The various uses could be quickly clarified in the methods section as there are only 3 things to test
 - Distribution similar?
 - Difference between categories of response (e.g. yes, no)
 - Difference between group (e.g. male/female) within a single category.
- It is suggested that it is just a case of being more certain (e.g. is the 67.4% on L348 an observation, or just randomness in sampling?) If two options, well informed or not, chance would be 50%, and with >100 respondents this is very probably significant for 10 respondent it wouldn't be.
- However, chi-squared for the whole distribution is fine where it shows something e.g. difference between genders for responses.

Thanks, and best wishes,

Dr Katharine Welsh, Associate Editor for *Geoscience Communication*.