

Response to Comments from Editor

Concerning the review of the manuscript “Towards family-friendly conferences in the geosciences: results from a first survey” by Elena Päßgen, Lisa Schielicke, and Leonie Esters submitted to Geoscience Communication (Preprint egosphere-2025-1200).

March 6, 2026

Dear Katharine Welsh,

We thank you for your comprehensive comments and detailed feedback.

In the following, the editor's comments are highlighted in bold and blue, and our responses are shown in black. Revisions concerning spelling corrections will be addressed directly in the manuscript. Our responses are structured according to the editor's review.

Minor revisions:

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

We added a reference to Blickenstaff, 2005 (<https://doi.org/10.1080/09540250500145072>), which addresses the leaky pipeline phenomenon in depth.

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

The sentence was changed to “*The underrepresentation of women is closely linked to the structural challenges they face such as few women researchers as role models and in visible positions, and a conflict between the structure and timing of academic careers with family-planning timelines (Welde and Laurensen, 2011).*”

Ln46-80 – Is there a 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.

The following sentence was added to the introduction: “*Although this issue is not specific to Europe and the US, the literature discussed here focuses primarily on these regions, as the survey analyzed in this study was disseminated primarily through European academic networks and German institutions.*”

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

The word “scientists” was changed to “academics”, as the data of Gottburgsen et al. (2022) refers to employees of higher education institutions.

Ln93 – Is this just relevant to scientific conferences?

“*Attending scientific conferences as a parent...*” was changed to “*Attending conferences as a parent...*” as our arguments and cited literature are not specific to scientific conferences.

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

We added a reference to Calisi et al. (2018) (<https://doi.org/10.1073/pnas.1803153115>), who argue that *“With solutions seemingly elusive, many women, and occasionally men, make a calculated decision to forego conference attendance and suffer the career consequences.”*

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.

We agree that it is important to acknowledge the bias in the data. We changed the sentence in the following way:

“Our study expands on the work of Pohl et al. (2022) by considering a broader sample that includes international researchers attending a variety of conferences in the geosciences, most of which are held in the Global North”.

Ln 158 – Remove ‘we’ and rephrase this sentence rephrased

The sentence was removed from the text as it was not needed to explain our statistical calculations.

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

The sentence was changed accordingly:

“When analyzing gender dynamics, we did not include non-binary participants and participants that did not specify their gender because their numbers were too small.”

Ln 312 – What does ‘not shown’ mean?

We agree that the phrase ‘not shown’ is misleading and removed it from the text.

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.

We agree that the claim of perceived loneliness is not sufficiently supported with the presented quote. Therefore, we removed this claim and changed the paragraph to *“Parents often reported feeling that caregiving responsibilities remain insufficiently normalized within the academic world. One participant expressed this sentiment with the comment, “Make it normal to have kids”.”*

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

The section was rewritten to clarify that the presented results were derived from the free text answers of parents (see below).

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.

The spelling errors are verbatim and we have indicated this with [sic] after the quote.

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.

Thank you for this helpful feedback. This section was completely rewritten, we included quotes to underline our claims and removed the suggestions. Some of them were included into Section 3.5 or the Discussion section.

Ln 667 – Where are these data from?

The presented data stems from interviews we conducted with conference organizers. We clarified this in the text with the following sentence: “ *The data presented in this section stem from interviews that were conducted with scientific conference organizers as part of this research.*”

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.

We thank the editor for their insightful suggestions on strengthening our statistical analysis. Following this advice, we sought guidance from an expert in statistical testing to finalize our analysis. We have incorporated a description of our statistical approach in the Methods section, as detailed below.

We note that for some analyses, especially those using only a single category and sub-groups, the number of participants was too small to perform meaningful statistical tests. We acknowledged this limitation in both the Methods and Discussion section (outlined below).

We emphasize that this study is primarily qualitative in nature, aiming to document lived experiences and foster dialogue between parents and conference organizers. The sample is neither statistically representative of the broader academic community nor large enough for comprehensive statistical analysis. To strengthen our narrative we will include the following sentence in the Methods section:

“Given the qualitative focus of this study, the survey responses are analyzed descriptively, with statistical tests used selectively to support specific observations rather than as the primary analytical tool.”.

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.

We added the note “ χ^2 test: $p > 0.05$ in all cases, distributions not statistically distinguishable.” to the captions of Figures 2, 3, 4, 6 and 8.

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.

Including only “yes”/“no” options yields the following result: $\chi^2 = 2.49$, $p = 0.114$. The distribution of yes/no answers is not significantly different between mothers and fathers. We included this test in our results:

“As shown in Fig. 4, 43% of the mothers and 24% of the fathers stated that they have already taken their children to conferences ($\chi^2 = 2.49$, $p = 0.114$ (non-significant); testing method A, other response options were excluded from the test).”

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.

We agree that greater clarity was needed regarding the specific purpose of each statistical test. The use of the statistical tests were clarified and numbered (A,B,C) in the Methods section, and we now reference the relevant method in parenthesis wherever a statistical result is reported in the text.

The following results have been supported with statistical tests:

“The need for support for families at conferences was particularly evident in the survey: 79.8% of the mothers and 57.4% of the fathers stated that they could not attend a conference due to family obligations in the past. Mothers agreed to this statement significantly more often than fathers ($\chi^2 = 6.93$, $p = 0.008$ (significant); testing method A).”

“With 92% of all fathers and 87% of all mothers, parents rely heavily on their partner when attending a conference (Fig. 6a., $\chi^2=19.02$, $p<0.001$ (significant); testing method B).”

“Fig. 6b. shows that parents who frequently attend conferences and workshops are significantly more likely to rely on relatives and friends for childcare than those who attend less often ($\chi^2 = 6.89$, $p= 0.032$; testing method C) .”

For the following observations, a meaningful chi-squared testing was not possible due to the low number of participants in each category:

“More of the parents that have already taken their children to conferences state that they visit three or more conferences a year (28%), compared to the colleagues that have not brought their children along (13%).”

“However, more parents with children older than six years reported that they have not brought their children to conferences (56%) and fewer of them expressed the desire to bring children in the future (23%), compared to the other two age groups.”

“Parents with children under three years rely on babysitters more frequently (33%) than those with older children (17%).”

As the results are presented as percentages, we have retained them in the text without statistical tests.

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.**

The following paragraph has been added to the methods section:

“The χ^2 test was used to identify the significance of our results according to these three questions:

- A. Is there a statistically significant difference in the overall distribution of responses between different groups (men/women, etc.)?*
- B. Is the distribution of responses across different response options statistically significant (e.g., was response 1 chosen significantly more often than response 2)?*
- C. Is it statistically significant that one response option was chosen more often by one group (e.g., men) than another (e.g., women)?”*

To acknowledge the limitation of the small sample size, we added the following statement:

“However, due to the small number of participants, we were only able to compile limited statistics. This applies above all to statistical tests in which the participants are divided into several groups (e.g., according to the age of the participant's children). Given the qualitative focus of this study, the survey responses are analyzed descriptively, with statistical tests used selectively to support specific observations rather than as the primary analytical tool. “

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.

We agree that it was important to confirm this finding with a statistical test and included the results of the chi-squared test to our observation:

"Our survey revealed that 67.4% of parents do not feel well-informed about support offers for families ($\chi^2 = 18.71$, $p < 0.001$ (significant); testing method B)."