Response to reviewers

Introduction: The Introduction is quite long compared with the rest of the paper, and I am glad to read that the authors intend to make it more focused. However, the Introduction (and the paper as a whole) lacks a clearly stated research question and a precise explanation of the study's purpose. Please make this clearer in the revised version.	The final paragraph of the Introduction articulates our research question and we have added text stating the purpose (to inform development of future rebuttals and revision of existing rebuttals). We have also added the research question to the abstract to reinforce the research goal.
Results: I found it very challenging to follow the Results section and understand the figures. Two of the reviewers have pointed out that the Results section was confusing in places. One issue is the complexity of the descriptions. I would like to challenge the authors to find clearer ways to describe the changes documented. Also, several variables are used in the text (CLES, B, t()) which are not defined. By looking at Figure 4, it looks like the change in accuracy decreases overall, with such a large increase in inaccurate responses. The figures show that the change in pre-rebuttal value of 1 changes positively in both a, b, and c. I understand this to mean that the "inaccuracy" of 1 actually increases. Shouldn't the change of the most inaccurate answers be negative? The text says that there was a "significant decrease in agreement with factual statements." Does this mean that the participants ended up agreeing less with the factual statements after reading the rebuttals on the website?	To add clarity, we've added the following sentence early in the Results "In this figure and throughout the results, we use a single measure accuracy where strong agreement with the factual statement and strong disagreement with the myth statement are designated the most accurate response." We've also added text throughout the results which we hope make the concept of accuracy more explicit and understandable. We also were more consistent with terminology, rather than interchanging between "accuracy" and "perceptions". We have added text defining CLES (Common Language Effect Size), B (the slope of the relationship between pre-rebuttal accuracy and change in accuracy), and t (the t-test statistic). Overall accuracy increases overall - factual accuracy decreases but myth accuracy increases more so the net result is an increase in accuracy.
Also, Figure 5c is very symmetrical. I read this to mean that roughly half the rebuttals led to a decrease in accuracy. In fact, most of the facts led to a decrease in accuracy according to Figure 5a. Have I misunderstood the figures here?	You interpret these figures correctly. Note that Figure 5 is a subset of our entire dataset, choosing just the 20 rebuttals that recorded at least 50 participants. Figure 5a (subset of the data) is consistent with Figure 4a (all the data), showing that overall, fact accuracy (e.g., agreement with fact) decreases. Conversely, 5b (and 4b) show that myth accuracy (e.g., disagreement with myth) increases. The two roughly cancel each out hence the symmetry in Figure 5c.
Discussion: With regard to the Discussion, I found the sudden introduction of the term "discernment" very confusing. We hadn't heard this term before, and then we are told that there was an "increase in discernment." What does this mean?	We have seeded the concept of discernment by adding a paragraph in the Introduction explaining it and why it's important.
Also, there is one point that is brought up in the Discussion and also in the Abstract. There, the authors say that the results indicate that "many visitors come to the site not to answer unresolved questions but to gather resources and answers." This is a bold statement, and I see no evidence of it in the data that was collected (also pointed out by Reviewer 2). The authors say in their response that this will be changed, as they cannot infer this "as strongly." I will need convincing that it can be inferred at all.	In this latest revision, speculation is fairly qualified now - starting with the qualified "One possible interpretation of this result is…" and added new qualifying text "(but we have collected no data to justify this interpretation)". We thought it important to include this concept because later when we talk about future research, we say "Questions specifically targeting motivations could address more definitively why readers visited SkS, better informing the website creators to meet readers' needs."
Other points: I was a bit surprised that the article does not reference the ongoing discussion in science communication about the deficit model. The Skeptical Science website can be seen as an example of a deficit model approach — one that seeks to change people's opinions by providing them with facts and information. There is, of course, nothing inherently wrong with this approach, and Skeptical Science is an excellent example of a community-based effort to counter misinformation about climate change. However, I think this article also speaks to the broader debate in science communication. The negative results presented here, in fact, align with the general shift away from deficit model initiatives and towards more dialogic and participatory approaches. As the authors note in their discussion, they are already considering new directions and developments for the website, which is very promising.	This is a good suggestion, we have added some discussion of the deficit model in the Discussion. We link it back to inoculation theory and in particular, logic-based inoculation.

The final major issue that needs to be resolved is that of ethical standards. The authors write that IP addresses were recorded. In most data privacy regulations, this is classified as personal information and therefore requires ethical approval before the research is carried out. This needs to be addressed in the article and explained in full. Please refer to Editorial: Geoscience communication – planning to make it publishable (https://gc.copernicus.org/articles/4/493/2021/) for my advice on this issue.	Our paper includes an Ethics Statement: This study was conducted with ethics approval obtained from the George Mason University Institutional Review Board (IRBNet number: 1379945-1). We also added text clarifying that IP addresses were deleted in the anonymised version of the dataset
Minor issues: Line 31: What are “stereotype attacks” that are referred to? The whole sentence is very challenging to understand. Please rephrase.	We’ve added some text to clarify what this means.
Line 61: What are “zombie arguments”?	We’ve removed this reference as it’s a bit jargony.
Line 117 (and the Abstract): What does “plurality of visitors” mean?	To be more precise, we’ve changed this to “nearly half”.
Acronyms: The paper flips between SkS and Skeptical Science. If you are going to use an acronym, please stick to it. In this case, the acronym SkS is not a well-known one and does not save much space. I would suggest keeping to the full name Skeptical Science throughout.	We’ve removed the acronym, only using “Skeptical Science”.
Line 212–218: I am unsure how the information in this paragraph fits with the arguments of the present study.	This paragraph is important to include because we are acknowledging that a limitation of Skeptical Science is that we’ve to date focused almost exclusively on science myths, not venturing much into solutions myths. And yet solutions misinformation is growing in prevalence (and in recent research I’ve found that it’s one of the most polarizing forms of climate misinformation). Hence an important development for Skeptical Science is the recent adoption of a large number of renewable rebuttals and future research will seek to test their effectiveness. We’ve restructured the paragraph to make this clearer. Note: the line numbers you quote seem to match the initial submitted manuscript, not the revised manuscript (which makes me wonder whether the comments on our speculation of why visitors came to the site is based on the initially submitted text rather than the revised, more qualified text.
Line 225: I do not understand how the analysis of the data in this study leads one to “guidance on ways that the rebuttals could be updated.”	We’ve clarified that the guidance is the inclusion of “replacement facts” that dislodge the myths being debunked.