



EGUsphere, referee comment RC1  
<https://doi.org/10.5194/egusphere-2022-244-RC1>, 2022  
© Author(s) 2022. This work is distributed under  
the Creative Commons Attribution 4.0 License.

## **Comment on egusphere-2022-244**

Martijn Kuller (Referee)

---

Referee comment on "Performance of the flood warning system in Germany in July 2021 – insights from affected residents" by Annegret H. Thieken et al., EGU sphere,  
<https://doi.org/10.5194/egusphere-2022-244-RC1>, 2022

---

This paper presents the results of a survey held among the residents of a recent severe flood event in Germany. It is a valuable piece of work both for the academic community interested in risk communication and for authorities to draw lessons from and improve practices in the future. While the overall quality of the work sufficient to warrant publication, there are a few issues that need to be addressed. Most importantly, the structure of the paper could benefit from clearer research questions, hypotheses and a methods and results section following this structure.

Overall the language is good and the paper reads well, but some phrases are rather "spoken" language than written language style (e.g., line 39: "save people's lives". I would rewrite to: "prevent fatalities" or "prevent loss of life". Second example in line 79: "...how the targeted population...". I would change to: "...warning perception by the target population..."). The article would benefit from review by a native English writer. Below I provide further details on this, as well as other comments by section, and then detailed comments in a table.

### Introduction

This section has a clear structure and explains the background, the need for the research and the research aim clearly. However, detailed research questions and hypotheses are missing, which are needed to justify the methods used (types of data (indicators) gathered, methods of data collection etc.).

### Data and Methods

It is stated that the survey was distributed via facebook, and a press release was sent through the newspaper. However, it seems like the survey could only be accessed through facebook. Why did the authors choose this method, and how did they make sure this didn't lead to a biased sample, as facebook is only used by a specific demographic, leaving out other specific demographics that might be important. Beyond age, general tech-savvyness and online presence play a role in the accessibility of the survey to people using facebook. If the authors somehow tried to correct for this bias, please provide some information about how this was done. If not, please discuss how the method might have influenced the outcome and what this means for the conclusions.

Also, it is explained what the questions in the survey were regarding, however I am missing some explanation as to why these questions were posed (what were the exact research questions the authors want to answer with the set of questions they selected, including the demographical questions), and what were they expecting to find? Part of this information could already be provided in the introduction. Furthermore, is the survey accessible to the reader somehow? It would be good to add them to supplementary materials for reference and future use by the readers.

The Likert scales used in the survey are analysed statistically, using the answers as numerical values. Although this is common practice, it is strictly speaking not valid, as these numbers are only category indicators, and not actual numerical values of the answer, and can thus not be used for statistical analysis (drawing mean, sd etc).

## Results and discussion

If the factors investigated and presented in Table 2 explain so little in the receipt of an official warning, what then is (or could be) the most important predictor for this?

All the results are self-reported data from the survey. I wonder how this self-reported data compares to behaviour and if this could be discussed in this section (e.g., how do the numbers of "knowing what to do" compare to the observed behaviour of people on the day?). Would there be any such comparison possible here in the discussion?

The results presented in Figure 5 show very little variation. This is a common problem with asking respondents what they want, choosing from a list of options: they are going to want it all. Such results are not very helpful when resources are scarce and trade-offs have to be made (in this case, we want to keep a warning clear and concise). A better way to measure such preferences could therefore be asking respondents to rank or to distribute a limited number of points among the options.

A section on further research is missing. I already made some recommendations in this review, e.g., comparing self-reported information to actual behaviour (if this is not going

to be implemented by authors).

## Conclusions

Because the research questions weren't completely clear (see other comments), the conclusion doesn't naturally follow from the rest of the paper. When this issue is resolved, the conclusion might be written in a more concise manner, going back to these research questions.

There are many references to the main text as well as other papers in the conclusion. I recommend rewriting the conclusion according to my above instructions, and avoid any references, as a conclusion should stand on its own, not repeat the results, but rather state the significance and meaning of results. As it stands, the conclusion mostly summarises the main text.

Furthermore, I think the paper could benefit greatly from a separate section (within conclusion or at the end of the discussion) that summarises the most important recommendations from this research. This research is very practical and applied, and has great potential to aid authorities around the world.

Please find detailed comments in the table below:

<b>Line</b>	<b>Comment</b>
39-40	One or two sentences discussing why the performance might have been so good for these events could benefit this section here.
153-158	This belongs more in the methods section, see previous comments.
159	Results are presented for which no methods have been described (logistic regression is mentioned). This should be added to the methods section (also see previous

comments about methods).

160

Please explain how to interpret odds ratios.

Table 2 and 3

From the table it isn't clear how gender impacts the prediction on knowing what to do. Is it male or female that increases the knowing what to do (explain in the table).

184

Is there data available to make a more precise comparison between the warnings issued by the different authorities and the warnings received by the population as reported through the survey? This could shed light on the effectiveness of the dissemination by the authorities. (a hint to this information is given in the text in lines 183-184).

189

The authors state that they find the trust in the credibility high. I would argue that 9% of people not believing a warning is quite a high number in an emergency situation, and less than half of the people stating they find the warning highly credible is to my perception quite low. From a government-issued warning, I would strive for credibility numbers close to 100%.

198

This sentence is confusing, please re-write.

210

This is more a concluding remark.

Fig. 5 and 6

Meaning of the numerical scale points need to be explained in the caption. Furthermore,

as indicated before, it is strictly speaking not correct analysing and presenting this data in this way.

298

It is not common to refer to tables and figures in the conclusion if not strictly necessary.