07 October 2021

IJERPH (ISSN 1660-4601)

Dear Editors,

Thank you for the opportunity to revise our manuscript “Visualizing the Invisible: Visual-Based Design and Efficacy in Air Quality Messaging”, Manuscript ID: ijerph-1363796. We are appreciative of the feedback provided by the reviewers and have attempted to address their concerns. We have included our updated manuscript and a point-by-point response to the reviewers’ comments below. Comments are in **blue**, and our response is in **red**.

Again, thank you for the opportunity to revise the manuscript. We believe this manuscript is much improved and hope you agree.

Sincerely,

Zoey Rosen

(on behalf of the authors)

Reviewer 1:

Overall Comments:

This study experimentally examined the effect of efficacy and visual design for messaging for air quality in the student environment. The paper is well structured and easy to read. However, the abstract must be improved. It should better summarize the main massage of the paper. The study analyses a rather small sample. Can we draw mayor conclusions based on 95 questioners? Moreover, the authors are using parametrical statistical tests on a non-continuous dataset. In this case, non-parametrical tests should be applied. I recommend revising the results and resubmitting the paper.

Additional comments are in the PDF document..

Thank you for taking the time to provide feedback! The abstract has been updated to be more specific to improve its overall message. We also recognize that the study does have a smaller sample, however, we believe that we do have a large enough sample size to conduct the statistical tests necessary for our results/discussion. The discussion has also been updated to elaborate more on the limitations in the dataset due to this particular sample. As for whether non-parametric tests should be applied, parametric statistics are robust with respect to violations of the assumptions of ordinal data and normality. We align with Norman (2010)'s assessment (citation provided in last comment in this response) that parametric statistics can be used with Likert data, with small sample sizes, with unequal variances, and with non-normal distributions. While we appreciate your recommendation, we will not be revising the results with different tests.

Specific Comments from Document:

[Line 10]: Which variables?

The abstract has been updated to read: This experiment assessed the effects of **message design and efficacy language** on students' visual comprehension, source 10 credibility, self-efficacy, and protective behavioral intention. [Line 10]

[Line 11]: Which hyptothesis?

The abstract has been updated to read "Hypotheses 1 and 2" [Lines 11-12].

[Line 13]: undefined...you have to be more exact in the abstract

Thank you for this recommendation. The sentence has been updated to read "...some statistically significant effects for efficacy and message design on **students' comprehension and protective behavioral intention.**" [Line 13]

[Line 138]: this speculation belongs to the discussion

This line has been moved to the discussion [Lines 452-454].

[Line 309]: this section should follow section 3.1

This section has been moved up and is now the new section 3.2. [Lines 273-288].

[Line 358]: did your data meet all the criteria for a parametric statistical test? You are not operateing with continious variables.

Yes, our data does meet the criteria for a parametric statistical test. Parametric tests can be used with ordinal data, commonly used in Likert scales, because the statistical tests are robust. The full argument is succinctly described using statistical theory in Norman (2010): Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Advances in health sciences education*, 15(5), 625-632. DOI 10.1007/s10459-010-9222-y

……

Reviewer 2:

Overall Comments:

This paper is important given the current concerns with wildfires and air contamination in the West Coast of the US. It is well-designed and written up. There are some issues that need to be clarified before the paper can be published:

1) The statements in the first paragraph of the Introduction need to be supported by references.

Thank you for your review for this paper. The introduction has been updated with additional references to further emphasize the support for the claims opening this study.

2) At the end of Section 2.1.2, the term "content self-efficacy" needs to be defined.

The section has been updated to include: "**Content self-efficacy describes the agency deriving from the knowledge that pertains to a particular situation; for example, content self-efficacy has been measured using social cognitive theory and earthquake preparedness literature to more specifically describe efficacy in an earthquake-specific context.**" [Lines 141-144].

3) In Section 2.1.3, it states that, "Self-efficacy describes one's evaluation of their ability to undertake a protective action." This definition is not an accurate definition of self-efficacy, which is defined as one's belief in one's ability to produce certain outcomes based on one's confidence in one's abilities. Is there a term such as "protective self-efficacy" to describe this type of self-efficacy?

Thank you for pointing out this discrepancy between definitions. The section has been updated to read "**From the risk perception scholarship that centered on self-protective behaviors, Kievik and Gutteling (2011)** describe self-efficacy as one's evaluation of their ability to undertake a protective action (e.g., "Am I able to protect myself from this risk successfully?"),

while response efficacy describes one's evaluation of the advised actions (e.g., "Is the advice I'm given about how to protect myself from this risk going to be successful?"). [Lines 149-154]

4) In Section 3.2.1, the content of the Air Quality Index should be explained.

The following information has been added to Section 3.3.1 (see previous reviewer's comments for why section was renumbered) explaining the AQI: "**The EPA's Air Quality Index (AQI) information provided in the message included an explanation of the numerical value that the index displays, such as what a value of 100 means. Additionally, information about what populations are affected by the different numerical values in the index was written to further explain the potential impacts of poor air quality and those most at risk for each designation on the index.**" [Lines 298-303].

5) The information in the second paragraph of Section 3.4 seems redundant with that which is presented at the end of the Results section. This information should only be presented in the Results section since it answers the authors' posed research question.

We agreed the information was redundant being in two sections. The second paragraph with the health-specific descriptive data has been moved to the Results section with the posed research question. [Lines 407-429].

6) The first three paragraphs of the Discussion section describe how Hypotheses 1 and 2 were partially supported and how Hypothesis 3 was not supported. Each paragraph should contain a more thorough explanation of the non-significant findings than simply small sample size, and there should be rationales given for each non-significant result.

Paragraphs in the discussion section have been expanded upon to have more thorough explanations of all of the findings, statistically significant and nonsignificant. More rationales are provided as well.

7) The discussion in the Discussion section of the National Weather Service being the most used source for air quality information was confusing. The authors state that seeking scientific experts and public health authorities online was encouraging; the National Weather Service does not seem to fit into either of those two categories.

The discussion section about the National Weather Service finding has been updated to read: "These findings somewhat support previous literature on the health information-seeking behaviors of college students which suggests that students are most likely to often or always use the internet for health information, followed by healthcare providers, and friends or family. **In the air quality context, students reported similar patterns in information-seeking; digital sites like the National Weather Service website were used the most often, followed by public health authorities, and then friends and family. Seeking information from experts in online spaces and in their personal networks within the domain of air quality is an encouraging finding. By referencing credible sources, such as the NWS or state and local agencies, students are displaying where they believe they can access information about the different hazards they experience.** Gaining a better understanding of where students go for information about air quality can be used by universities in order to boost their message credibility around the health risks during wildfire season, and can be important for **community stakeholders in**

preparation for mitigation messaging, building community partnerships, and other pre-season strategies." [Lines 502-515]

8) The authors did not mention the limitation of all students being from the Department of Journalism and Media Communication. There are generalizability limitations from only using one college department. Also, students in this major may be particularly attuned to how messages are transmitted or used to interpreting media messages in a wide variety of formats, which could account for the lack of significant differences. This limitation should be mentioned.

Information about the department has been added to the limitations section. Students recruited through SONA have to be enrolled in a journalism class, which does speak towards the generalizability concerns. However, a majority of the students in the sample pool were enrolled in a common-core technical writing course, which includes majors from the entire university. This detail will be added as well. [Lines 528-539].

9) On the last page in the Conclusion, the authors state that there were statistically significant effects on comprehension, source credibility, and protective behavioral intention. However, there were no statistically significant effects on source credibility.

Thank you for catching this discrepancy, the conclusion has been updated to state that there were statistically significant effects on comprehension and protective behavioral intention [Lines 552-554].