Reviewer’s report

Title: Languages for different health information readers: Multitrait-multimethod content analysis of Cochrane systematic reviews textual summary formats

Version: 0 Date: 15 Nov 2018

Reviewer: William Speier

Reviewer's report:

The authors present an analysis of the readability, tone, and style of different formats of scientific text from the Cochrane web site in several different languages. The topic of readability is vital for the understanding of medical literature and has been shown to affect medical decision making. In particular, I appreciate the inclusion of translations of the clinical text as most readability assessments that I have seen focus on a single language (usually English). The analyses performed by the authors are conducted well; the results are interesting and generally agree with previous studies that clinical and scientific text, including "plain language" summaries, is generally written at too high a level for the general population to understand. While I believe the importance of the manuscript is high and the analyses are appropriate, there are a few areas that I believe are currently lacking which should be addressed.

1. The background section should generally be expanded. While the authors include some good references of previous studies that have tested readability of health literature, the background does not currently mention these studies explicitly or how the current study builds off of previous findings. The authors also do not fully explain what Cochrane is and the reasoning behind using it. For context, it would be helpful to have a brief review of some other efforts to make health information available to the public, describe how Cochrane compares to the other efforts, and say why it was chosen for the current analysis. Mentioning who performed the translations would also be helpful information as that can affect the readability and accuracy of the translations. Finally, the formula for the SMOG index is not fully explained and the authors do not discuss the other methods for readability (e.g., Flesch-Kincaid) and why SMOG was chosen for this analysis.

2. The results are interesting, but the presentation is inconsistent and can be hard to follow. For instance, the sentiment analysis section is the only section that includes numerical results while the other sections use qualitative statements such as "did not contain strong emotional tones". The figures also do not convey the variance for the values, which can make it difficult to appreciate the differences between the different document types. Box plot might be a better way to demonstrate these distributions.

3. Finally, the authors do a good job of presenting analyses, but the hypotheses behind the analyses are not clearly stated. For instance, why is it important that sadness is the
dominant emotional tone? Does that affect a reader's ability to understand or interpret the text? What is the practical significance of different languages having different reading levels? The manuscript would be stronger if it had a firmer conclusion and description of how the results of this study could be applied to improve the readability problem.

Other comments:

1. It is interesting that Croatian and German have much lower SMOG indices than French or English. Is this because of the translations? Is there something specific about the languages that effects the readability scores?

2. It is interesting that scientific abstracts have significantly lower analytical tone (figure 3) but significantly higher values for analytical thinking (figure 5). What is the difference between these two (other than the analyzer used) and why might it have caused this apparent discrepancy?

3. In general, statements about "high readability" can be confusing because "high readability" means "low readability scores" due to the concept of readability being inversely related to the SMOG score. I think the authors even mix things up in at least one place ("decreasing readability" on page 8). I would suggest saying "more" or "less readable" when comparing two types of documents and "high SMOG grade" (or "high education requirement") when talking about absolute readability.

4. Some acronyms are not defined (e.g., SMOG). Also, Linguistic Inquiry and Word Count is defined twice, first as (LWIC) and then as (LIWC). Both versions are used multiple times (LIWC is correct).

5. The authors make a statement about "truthful information about health", but I think they mean accuracy or precision. Truthfulness usually includes connotations about honesty. I do not think the authors mean to say that changing readability will lead to deception, but rather that simplifying language may necessitate some loss of precision.
Are the methods appropriate and well described?
If not, please specify what is required in your comments to the authors.

Yes

Does the work include the necessary controls?
If not, please specify which controls are required in your comments to the authors.

Yes

Are the conclusions drawn adequately supported by the data shown?
If not, please explain in your comments to the authors.

Yes

Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I am able to assess the statistics

Quality of written English
Please indicate the quality of language in the manuscript:

Needs some language corrections before being published

Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?
If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal