**Overall Comments**

As I am not intimately familiar with the EU funding proposal procedure, it was interesting to read the manuscript and delve into the intricacies of the process. The authors also used analytic methods that were novel for me and I enjoyed learning about these methods.

However, I have a number of concerns about the manuscript, many of which I addressed directly as comments on specific sentences or areas. Overall, the structure of the manuscript needs some work. Most scientific journals, PLOS One included, follow this structure: Title page, Abstract, Introduction, Methods, Results, Discussion, Conclusions, then Acknowledgements, References, etc. This manuscript includes a Background section, the Results and Discussion section have been somewhat combined. The Conclusions section reads partly as a Discussion section.

The manuscript is also quite difficult and heavy to digest and I found many instances of redundancy. The authors should carefully revise each sentence to ensure they are not repeating something that has already been said and ensure that their manuscript is the most concise representation of their work as possible. This is important to guide the reader, make their overall points as clear as possible and promote success in publication acceptance, as 17 pages is far too lengthy.

Finally, all sections could be shortened to only report what is necessary. For example, the whole section in the background on the EU funding process is presented in 5 paragraphs. This could be broken down into a single paragraph, perhaps two. I highlighted a section that I especially find unnecessary, but the whole section could be shortened.

I have included more specific comments below, organized by sections.

**Background/Introduction**

I’m interested to know who the evaluators are for EU proposals and how they are chosen. How many evaluators are there for each proposal and how do they decide which proposals to accept if there are two with similar scores? As the evaluation process is the critical period that determines the outcome, it would be insightful to provide more information on this aspect.

How many projects were funded by the H2020?

**Methods**

The authors state that 16 proposals were sampled, but these were all proposals that their institution was directly involved in. This would likely induce a significant sampling bias that is not addressed, as these proposals would most likely not be representative of all proposals in the H2020. What proportion of proposals in the H2020 are represented by this selection? What kind of distribution did the proposals have across themes/subjects?

Why did the authors limit each call to 2 projects? How many accepted/rejected proposals did they have per call?

In Section 3.1, I suggest making a flow chart because it is difficult to follow the numbers.

**Results/Discussion**

The results section should be separate from the discussion. There are instances where the authors compare their findings to those of other studies. This comparison should be strictly reserved for the discussion and should not be present in the results.

The results section could be broken down into sections for ease of reading. It should be more concise, with the findings of the study clearly stated without interpretation (interpretation of results is reserved for the discussion). However, the findings in this study can be compared among each other (ie. Compare table 5 with the parsimonious solutions to Table 6 with the intermediate solutions). Again, these comparisons should be without interpretation, only facts and findings should be stated without the author’s opinion on what these findings suggest.

There are too many results tables in the manuscript. The authors must choose the most pertinent tables to present and move the rest to the supplementary material.

The authors can also try to find a way to condense the findings, rather than just move a number of the tables to the supplementary material. For example, Table 9 is not necessary to present in the main results, or even at all, as the most important information is summarized in the text (the range of hamming distances, mean, median). It is not necessary for the reader to see each hamming distance for each pair of proposals when there are summary statistics available. The same could apply to the other tables; the reader does not necessarily need to see each raw or unique coverage value. Present the most important findings in the text and move the rest to supplementary material, summarized when possible. Also, numbers should be rounded to 1 or 2 decimal points.

**Conclusions**

The conclusion section reads partly as a discussion and partly as a conclusion. The discussion and conclusion sections must be better distinguished. The first paragraph of the discussion should ideally be a sentence or two about the importance of the study/what this study adds to the existing knowledge, followed by a brief (2-3 sentences) summary of the main results. The conclusion section should not be very long, a brief paragraph no more than 5-6 sentences to summarize the findings of the study, and perhaps again accentuate on the importance of the study.

The Conclusions section (which to me is partly the discussion section) is very difficult to read. The main findings of the study are not clearly presented and it feels like the authors do not directly discuss the factors influencing proposal acceptance, as it is often quite vague and often there are not specific conditions mentioned. A discussion section should remind readers of the main results found, compare the result to those of other studies, then summarize what the finding suggests overall, given all the evidence. The authors should rework the discussion section to more concisely state their findings and remove all redundant sentences.

The authors included a section about the limitations of their study, but not about the strengths of their study. It is important to state what makes this study strong. Consider blending the strengths with the limitations or minimizing the limitations. For example, “the binary encoding method simplifies the complex characteristics… but, by using a binary coding method we were able to…”.

**Suitability for PLOS One**

The manuscript is not in the proper format for submission at PLOS One. The authors must examine all formatting requirements on the PLOS One website. The structure, as mentioned previously, is not in accordance, as PLOS One requires a Title Page, Abstract, Introduction, Methods, Results, Discussion, Conclusion (optional), Acknowledgements, and Reference section. The basic formatting requirements must be considered (text should be double-spaced, page numbers should be included, continuous line numbers should be included). The authors must also pay attention to the formatting of their references, which is slightly different from the formatting for PLOS One. For example, after six authors then *et al*. should be employed, there should not be brackets around the year of publication, etc.

There are many other specific formatting requirements that must be considered. I encourage the authors to follow the journal’s formatting requirements on their website: <https://journals.plos.org/plosone/s/submission-guidelines#loc-reference-style>