# Response Letter: Talk or Not to Talk: Factors that Influence Communication around Changesets

We would like to thank all the reviewers for their insightful feedback and the effort they put in to improve our paper. We believe that the latest version of our paper corrects the issues pointed out by our reviewers. We will now address each point made by them, starting with the three required improvements by the Coordinator.

## Clarify about what kind of systems the study is trying to address: recommender systems or communication support systems? Please elaborate about this point.

The study addresses recommender systems. At several points in our original submission we used the term “communication support systems” to highlight the information-seeking aspect of recommender systems. We removed such mentions for the current version, and replaced them by “recommender systems” as appropriate.

* *Is the information needed something that can be generated/extracted/mined from existing data or has to be solicited from developers’ interactions? Furthermore, if it is for designing recommender, more likely from what I can infer, information to be recommended not be what the developers would explicitly seek for or consider, but rather what they are not aware of. Therefore, less likely to seek information may not indicate that systems should not recommend, but just developers do not have specific needs in mind.*

The information that recommender systems would need can be extracted in part from existing data, and in part has to be solicited from members of the organization. However, we did not investigate which of the guidelines we proposed require eliciting information from developers and which do not, as we judged that such questions were out of scope for our original study.

## Improve the data analysis and the report of the results.

We made two major changes to the representation of the data analysis: (1) Instead of reporting the reverse rank sum we report the average ranking and (2) we included the ranking distribution of the different items using sparklines.

* *How does the weighted item ranking for the survey affect the results?*

*The authors ranked survey items by aggregating points each participant assigned. This simple aggregation does not normalize the points or indicate individual differences. For example, item A with 14 points plus 6 points will be seen as the same to items B with 10 points plus 10 points. B is consistently considered highly likely to have influence but A is controversial.*

To simplify the ranking we now use the average rank of each item, instead of the sum of reverse rankings. Furthermore, to dispel doubts about large disagreements among our participants, we included a table that shows the ranking distribution of each item with a sparkline. The table shows that there is a tendency for agreement among developers. The agreement is strongest for the most extreme cases (the first few and the last few of each subsection), which are the focus of our discussion.

* *The demographics of survey participants are not included in the paper. What are their expertise levels, roles, locations and so forth? Will this composition introduce any bias into the conclusion?*

We included the information about which continents the developers are from as well as their roles. Furthermore the paper now acknowledges the potential influence of the cultural differences in our results, especially from the interviews and observations.

* *The results shown in Table 1 appear to be survey-centric. It is not clear whether issues or factors other than survey items emerged from interviews and observation. Nor is it explained whether any discrepancy arose between data resources.*

Table 1 is survey-centric. The results reported on the paper, however, are not: some of the factors we discuss did not arise from the survey (for example, the last finding in the current version describing the discrepancies between the pride of developers in their work and the actual business goals).

The discrepancies we encountered were minor and they do not affect the findings we report. Therefore, due to space constraints, they are not mentioned in the paper.

## Consider to connect the design guidelines with the findings more tightly.

* *One improvement the authors may want is to connect the design guidelines with the findings more tightly.*

We agree that this is an important improvement. We adjusted the presentation of the guidelines by first stating what we learned from which finding, followed by the implications, and finishing with your recommendation.

* *The authors could regroup all the 7 findings in the similar way as the 4 guidelines are organized.*

We regrouped them so that the findings align with the guidelines. We also changed the order of the sub-tables to better reflect the order of the findings.

## Discussion of the threats to validity needs to include the cultural influence.

We added an acknowledgement pointing out the potential threat of bias due to cultural influences. Furthermore, we now mention that we intend to follow-up with interviews of developers from different sites in the next stages of this work, to mitigate the risk of cultural bias.