We thank the reviewers for their comments, and think that all of the suggestions can be addressed by minor clarifications, specifically:

The primary goal of our study was to *relatively* compare potential risks with each other. Our study 1) relatively ranks a wide range of potential scenarios, 2) confirms that privacy outnumbered all other concerns, and 3) shows users’ other concerns. Regarding B’s concerns about our claims: we performed the first study on wearable device users’ risk perceptions to involve over 100 subjects (or in our case, almost 2,000). From the papers cited: neither Kapadia nor Christin performed experiments, and instead examined existing literature; Klasnja interviewed 24 participants from a single city; and Raij interviewed 66 participants from a single university. We are happy to cite these papers, but believe our study examined a much broader set of risks across a much broader sample--therefore contributing to the existing literature.

To A’s design questions: we used a mixed factorial design for our IRB-approved user study. To B and D’s methodology concerns: we intentionally did not discuss utility because we were only interested in relative risk perceptions. If participants were biased, they were biased equally across all treatments; therefore the relative differences observed are still valid. As we note in the beginning of 2.2, the order of the survey sections was randomized.

We have a rich dataset of open-ended responses, ranging from 1-word to 143-word responses. Taking B’s suggestion, we will clarify our coding labels for the open-ended responses. We are in agreement with A and will include more on these open-ended responses.

We agree with A’s suggestion to merge Section 5 with 2.1. However, we disagree that factor analysis should have been performed: factor analysis is not appropriate for a mixed factorial design. (If each participant received the exact same questions, factor analysis would have been appropriate.) Instead, we used multiple independent coders, which is a widely accepted method for categorizing data. Regarding A’s questions about our regression model: we found no significant interaction effect between data and recipient or with wearable ownership, which is why we did not include them in the models. We can explicitly state this in the paper. Finally, based on our focus group feedback, we shortened our survey so as to prevent participant fatigue, which why we gave participants 2 of the 5 smartphone questions, rather than all of them. We are happy to state this, as well as the average completion time (11.5 minutes) and responses per question (141 per data type, 744 per smartphone risk, 2,779 per recipient, and 35 responses per each unique data type/recipient combination).

We agree with C that bystanders are an interesting angle of research; however, our study focused on the user of the wearable device. We did this because we felt that the user of the wearable device would be most at risk, so we wanted to isolate this consideration. We agree that additional work is needed to examine bystander perceptions, but we believe that this is orthogonal.