Thanks to all reviewers for offering so many valuable suggestions! We have taken all the suggestions and made significant revisions to the paper. We describe those below:

1. *Clarify research contribution and generalized benefits to other CSCW practitioners and researchers*

The paper reports specifically reports on a collaborative information analysis task situated in the intelligence domain, but our findings regarding team process and breakdowns address the interests of the broader CSCW community. In the revision, we more clearly emphasize, in both the introduction and discussion, the contribution of this paper to understanding and suppoting the coordination of complex collaborative work: (1) we documented a spontaneously adopted interleaving workflow involving switching between modeling and analysis, and quantitatively characterized how an earlier switch from modeling to analysis improves team performance; (2) we distinguished three labor division strategies and five factors that impacted team performance; 3) our result implied that effective collaboration requires teams being aware of not only partner actions, but also contribution value, uncertainty, and context of insight.

1. *Collaborative modeling and annotation literature*

We developed our treatment of prior literature in collaborative modeling and annotation.

Many researches [@chen2014, @Nokelainen2015] have demonstrated that annotation helps share knowledge when reading and writing, two critical activities in intelligence analysis.

Prilla et al. [@Prilla2013] gave a comprehensive review of collaborative modeling.

Comment: I don’t think we want to imply that the revision was just “more” literature review. That does not seem to be a strong rationale.

Comment: I don’t think we explained here HOW the literature review was developed or HOW that improved the paper?

1. *Comparing use of CAnalytics to other existing tools*

The research setting and research design we adopted has strengths and limitations. A key strength for us is that we were able to work with motivated participants on a pretty complex collaborative task through a two-week period. Because our participants were studying our task domain, they were all aware of standard tool approaches such as ACH and Analysts Workbench. These are important strengths, particularly in our domain where expert practitioners are difficult to find at all (many work for intelligence agencies and contractors) and their time and effort is closely managed.

One of the limitations in our study design is that we were unable conduct control group comparisons. Ethically it is difficult to assign students to education conditions that may be disadvantaged, indeed, the instructor we worked with wanted all of his students to experience the same educational opportunities. This is a direct conflict between our interest in using the classroom context as a larger-scale testbed and the students/instructor interest in eing and learning aboutWe have strengthened and clarified our explanation of strengths and limitations of the research design. The classroom is surely a special case of the “real world”, but it is the real world relative to a lab study context. We often cannot run control conditions in workplaces.

This limitation also bears on the issue about student groups “opting out” of using our tool. Students often do not follow the recommended instructions; this is just a fact about the classroom context. Moreover, from our ethical standpoint as researchers, our participants always have the free choice to not participate at any time. Therefore, this issue was beyond our control. We reported the fact that one group did not participate.

Comment: We used the term classroom study in our title but perhaps underanalyzed what a classroom study is and the strengths and limitations of doing one.

1. *Add description the use of CAnalytics outside class, what other systems they used, and the effect*

We added data analysis in the beginning of result:

Teams had three intensive usage sessions over the week, although they could access the tool any time; two sessions were in class and one was outside class before the team report deadline. 22 teams self disclosed that they used CAnalytics as the analytic tool in the project although they were allowed to use any other tool; one team reported that they mostly used Google Doc. The reported usage was confirmed by the system log. Seven teams reported using GroupMe and other instant message outside class. They used these tools for instant communication and coordination of meeting. Ten teams reported using Google Doc. Nine of them used Google Doc only for composing the final team report and CAnalytics for analysis tasks; one team went further and used Google Doc as the main analytic tool.