Please see our responses to each reviewer comment below in BOLD, below the corresponding comment from the reviewer.

General comments:

Interesting article about enhancing the T-shaped learning profile of hydrology students. In the article a comparison of a DMDGC simulation module with a paper laboratory module. It is hypothesized that students who followed the DMDGC module would demonstrate a better understanding of theoretical and applied hydrology concepts related to flooding in a contextualized and realistic scenario and that the simulation condition would lead to a better understanding of teh professional role of hydrologists. The DMDGC model produces a visualization of modeled and observed hydrograph results. In the paper module students had to perform hand calculations. In fact it is a comparison between a traditional paper pencil method with a computer simulation method, asking whether the latter method is more effective than the first one. It is good to read that the use of a simulation model can enhance student's knowledge and understanding of the hydrology field.

We thank the reviewer for their positive comments regarding the manuscript.

Specific comments:

About the methodology, it is not clear why the group sizes of the two groups differ. Why does the DMDGC group consist of 52 students and the control group of 36? What criteria have been used to create this difference? As far as I can read also no further analysis took place on students' backgrounds and preferred learning styles which might have influenced the outcomes of this study. Also no information is given on the results of the pretest. Were the 52 DMDGC students better than 36 control students. How did the allocation of students to each of the two methods take place? Has this allocation influenced the result of the investigation?

As this investigation represented an implementation in an actual classroom section at a local community college, the difference in the number of participants across the groups was a natural reflection of course enrollment, by lab section. This factor was outside our control, however we would caution the interpretation that either condition contains, in any way, a small or inadequate sample size. Further, issues of unequal sample size are often founded on the concern that such small samples would in fact bias the sampling of critical variables by creating unequal variances between groups, and thus undermining the estimation of the experimental effect. However, statistically speaking, this was not the case in the current experiment, as Levene's tests for all included comparisons produced a non-significant result ($p > .05$) across the group variable, confirming that the variance across groups was equivalent despite the difference in sample size. In other words, variances across the groups, and for all measures, were statistically equivalent as measured here. Coupled with the robustness of the ANOVA/ANCOVA procedure relative to violations of the assumption of equal variances, we are confident that this small disparity in group sizes did not affect the evaluation of the

current manipulation. We do appreciate the concern however, and will add a reference to this fact in the revised manuscript.

Related to this issue, all students in the current study were also drawn from the same population of students (e.g., community college) that self-selected to enroll in this course (without the knowledge that this experiment would be part of the curriculum). As such, while numerous demographic differences were not explicitly evaluated, it is reasonable to expect that these students are more or less equivalent on educational background, SES, etc. We would also caution any consideration of learning styles as a relevant variable, as there has been much research in the field dismissing such assertions as incorrect (Pashler et al., 2008).

Finally, related to the above points, we would also like to emphasize that we are comparing differences in learning within participants, albeit across groups. If one were to concede that the participants were in fact different in each group (which again we would not), these differences are in fact controlled for in a gross sense as we are evaluating the participants progress against themselves. Perhaps the most critical variable that might affect the accurate assessment of these knowledge gains (e.g., prior knowledge assessed via the Pretest), was also in fact explicitly controlled for by the ANCOVA procedure, by utilizing Pretest performance as a covariate. Group means as presented in Table 3 represent adjusted means relative to this covariate, thus again controlling for any differences on the pretest. The reviewer is correct, in one sense, that it is always possible that other demographic variables might interact with this change, however we would suggest that this should be a topic for future research. We would also caution that the explicit control of such other variables (the reviewer does not explicitly identify specific characteristics) would also reduce the ecological validity of such investigations, which we see as a critical contribution of the current work.

T-shaped learning profile. Perhaps it is my lack of knowledge and understanding about the DMDGC module, but it is unclear to me how this module, has enhanced with the students the understanding of the role of hydrologists. It is said that the lectures, which were content wise the same for both groups, focused a.o. on the roles and responsibilities of agencies that provide flood prediction and management services in the USA. How has the simulation model helped to improve student's understanding the professional role of hydrologists?

The reviewer is correct that all participants received some consideration of the role of hydrologists in the lecture component of the course, however, it is our contention that the students in the DMDGC condition gained a better sense of this professional role by actively engaging in the DMDGC exercise. So in other words, rather than understanding the role of hydrologists in an abstract sense (likely conveyed via lecture), students who interacted with the DMDGC received a better sense of the day-to-day duties (meaning job skills) that hydrologists practice. Thus, it is our contention that the DMDGC exercise represents a realistic approximation of job duties of a hydrologist, whereas the paper and pencil lab sections still convey this understanding in a less explicit and more abstract way, as evidenced by the increase in appreciation of these duties by the DMDGC group.

Secondly, T-shape learning should not only focus only on widening one's own field of expertise; i.e. focusing on the professional role of hydrologists. In daily practice professionals should also be able to speak to people from other domains. Students should also be trained in this respect. So, this study is limited in its scope. About the learning outcomes. These are very poorly formulated as they do not say anything about the level of knowledge and skills students . Blooms taxonomy is fully lacking in this respect. The outcomes as they are described as such do not say anything about how well and at what level students have mastered these. Have the students been informed about these outcomes before the start of the course?

We agree wholeheartedly that part of the job duties of any hydrologist (or even more broadly, scientist) is to interface with other individuals, both across fields and outside of the field (e.g., the public) effectively. However, such training is outside of the current scope of this experiment, as (to use the reviewer's own suggestion), the manipulation utilized here is designed primarily to address the cognitive aspects of Bloom's taxonomy. Not only are we evaluating the gain of knowledge in current areas (e.g., identified by learning outcomes), but we are also evaluating the application of these knowledge states (i.e., expert ratings of effectiveness). As such, we would argue that Bloom's taxonomy is alive and well within the current experiment, although we would simultaneously add that Blooms's taxonomy is only 1 of many potential means to defining learning outcomes. We would also finally add that the learning outcomes identified here are consistent with major learning outcomes across the field of hydrology, and thus it is important to evaluate learning interventions within the context of said outcomes, again to promote external validity.

To the reviewer's final point, students were not made explicitly aware of the learning outcomes in their final form, as this could potentially bias student performance while learning. We wished to minimize such influence in an effort to provide a better estimation of the experimental effect. It would be of interest, however, to evaluate whether the presentation of such learning outcomes might magnify the current effect, as there is much classic research in the fields of cognition and education that suggests that presented organization affects how individuals encode information. However, we again stress that this might be a fruitful area for future research, extending the current findings presented here.

Technical comments The reading of the text could be improved to include table 3 and figure 2 in the text.

We see the reviewers point, and would be happy to move these tables/figures should the editor deem it necessary. These materials currently appear at the end of the manuscript in an effort to remain consistent with APA style.