Thank you for your comments on our manuscript. We have made significant modification to the text to reflect your suggestions, and feel that the paper has been greatly improved. Below we have responded to each of your individual comments.

-The introduction gives a very good overview of the problem and state of the art. -The article claims to be studying ravines or permanent gullies. Yet this experiment is clearly tailored at measuring what happens during a newly forming incision, that erodes from scratch. I believe the results would be different if erosion would be analysed in an existing channel (or existing permanent gully). These might well respond to changes in discharge.

- Our results are focused on the development or growth of a large permanent gully or ravine where a steep knickpoint is propagating up onto a flat terrain, they would certainly be different for an existing channel.

-The choice and description of materials is crucial and should be better explained. "Mud" is not an objective description as far as I know. The 96 micrometer substrate classifies as sand following objective classification systems (0.05 – 2 mm) and the finer one as silt.

- The term "mud" has been replaced with "fine grained" throughout. The materials were selected based on what was available, yet they were effective at capturing two end members of a system.

What are the indications for cohesion and that the finer substrate really acts as detachment limited? What is the critical shear stress of these materials? Other properties such as angle of repose will also greatly influence wall stability and could be useful to include.

- The cohesive substrate was determined to be detachment limited by modeling the ravine growth using detachment limited equations. It is described early as detachment limited which is supported later in the results and discussion sections. In addition, the fine grained substrate forms steep vertical walls, which add qualitative support to the idea that the substrate is cohesive and is likely to behave as detachment limited. We agree that these other properties may be useful, but they were not measured in the experiments. In future work this should be considered.

-The experimental setup strikes me as odd to investigate ravine erosion/concentrated flow erosion. Why was a flat bed used? This way the flow does not concentrate until the outlet? From the only picture that is included of the experiments (Figure 1) it seems like you have multiple gully/ravine heads in the mud substrate. What is the effect of this on your results?

- The experiment was originally designed to examine the impacts of hydrology changes on flat agricultural landscapes. We believe that the experiments are interesting without representing a specific case study, so the field area that was the impetus for this study was left out. By only forcing the flow to concentrate at the outlet we are able to investigate the process of head cutting, which is a dominant mechanism for ravine growth.

-The presence of several knickpoints in the mud case seems to indicate the presence of plunge pools. These could potentially be very important yet I am missing an explanation on this. See for example results by Govers et al. Earth Science Reviews 2007.

- In some cases there was a small secondary gully head that formed, yet a typically there was a larger head that was the most dominant. These are an excellent reminder of the natural complexity in a system. Here while we tried to control the system closely, there was always quite a bit natural variability. We have included a brief description of plunge pools in our discussion. We have elements of plunge pools in our system, yet are missing other elements, such as significant undercutting and blocky failures.

-The presentation of experimental results is quite confusing. The authors use a mix of "water delivery rate", flow rate and discharge (figure 3), sediment volume removed versus mass (it should be easy to convert the first into solid discharge rates using the bulk density of the sediment) etc. Also a mixture of units is used (m3 in the text versus cm3 in figure 3)

- Discharge and m^3 is now being used throughout. The volume vs mass is a result of each graph being presented in the units the data were collected in. Bulk density was not measured for each experiment and therefore the conversion cannot be made.

-It is not surprising to me that no good relation can be found with discharge. Previous studies, mentioned in the introduction, clearly indicate that gullies grow fast in early stages and then reach a more stable state. Figure 4 and 5 clearly shows this as sediment flux peaks in the beginning and then declines.

- While other studies do find a rapid early channel evolution followed by a stable state, they do not explore how discharge impacts that period. We would argue that our data suggest that the length of the rapid growth phase may be dependent on the flow rate. It is also worth noting that it isn't clear from our data that all channels reached that stable state, only in our lowest discharge case did the channel erosion approach zero, in some experiments that sediment discharge continued to increase or was stable throughout the experiment. It might be interesting to extend the experiments to lower volumes of water, I am inclined to believe that the results would be similar.

-The text includes a lot of statements that are not backed up by data. For example, p.6 lines 11-12: "channels in the sand erode due to head cut propagation as well as lateral channel migration". No picture or DEM is included however comparing the results and evolution in both substrates so there is no way to check this. The only available figure 6 is difficult to compare and raises further questions. I recommend to align this graphs in the same direction (flow direction for example, with the outlet facing up). In the mud, two gully heads have formed. This confirms my point made earlier about the experimental setup. How was this handled in the data analysis? As flow splits, how was this modeled? Is the measured channel width the sum of both channels? could this explain why you see no increase in width with increase of discharge for the mud case. See papers by Torri et al. on using channel junctions of rills to model width downstream (10.1016/j.geomorph.2005.11.010).

- The image in figure six shows some lateral channel migration because the channel eroded along the entire length and was made wider, but did so only on one side. All images will be made available in a supplement which will make this clearer for other experiments as well so the statement is not simply an observation, but is rather backed by additional data.
- The graphs have been aligned the same direction.

- Channel width was measured downstream of the confluence between the multiple channels. This was clarified in the text. Channels did widen as discharge was increased, they simply didn't widen downstream of the newly eroded areas. We might expect that given enough time to fully adjust to the new discharge that the downstream reaches of the gully would have also widened.

Another example are the statements about slope that are made in the text, yet no discussion is given on the slope results presented in table 2. -In paragraph 3.2 the authors are mixing results and materials and methods or model description. I believe these equations should go in materials and methods.

- The methods section focuses on the methodology of the experiment to answer our guiding question for this research. Additional information is provided as analysis in the results and analysis section of the paper. As these equations focus on additional analysis rather than the methods of answering the overall guiding question they are not appropriate for that section.

-p8 line 5 "our sediments were easily eroded" Isn't this contradictory to claiming the mud case is detachment limited?

- Once detached the sediments were easily removed from the system

-What is the use of describing your experimental conditions in a table if you then cite flow rates etc. in every figure (example fig5)?

- The graphs were modified to use run name rather than indicate flow rate in the legend.

-p.9 lines 1-4. This is an interesting conclusion. I think it merits some expansion, which comes after line 25. I think this would be clearer if the two paragraphs that are in between were moved. This width-discharge relation is used commonly in gully erosion models but lacks differentiation according to material type as far as I know. Did you measure top width or bottom width? I could not deduct this from the materials and methods. In many field studies bottom width was measured.

- The paragraphs mentioned were shifted.
- Channel top width was measured, yet in our substrates bottom width is essentially equal to top width. In the fine grained substrate this was a result of the sediment's cohesion which formed nearly vertical walls. In the sand substrate this is a result of the banks being very low and therefore the bottom and top were impossible to differentiate.

-Why not use the more common term gully erosion in stead of ravine erosion? (2844 terms in WoS versus 212 results)

- We feel that the term "ravine" more accurately describes the features being studied rather than the more general term "gully".

- if SI units are not deemed illustrative, please use at least standard abbreviations throughout the text for example p5line 1: ml/s figure 2 time (min) not (m) -table 1. Why call this flow? Flow rate or discharge.

- Abbreviations have been standardized in figures and text.

-The quality of the figures is not very high and looks like they have been done using a standard spreadsheet. Please improve using a more professional graphing programme (R, sigmaplot, grapher, whatever,. . .).

- Thank you for your feedback

-p.6 lines 1-2 and assuming that you have Dunnean runoff generation as you have in your experiments. With Hortonian overland flow generation, the whole situation changes.

- Thank you

-p.5 line 28. Explain meaning of the threshold line. -p7 line 2 Nachtergaele -figure 7. Please indicate the exact values used for b. 0.4 or 0.39 (text??). What is the value in case a.

- The values reported for b in the text and on the figure (0.39 for the sand and 0.27 for the fine grained substrate) are the values that gave the best fit for our data. These are slightly lower than those found for gullies and ravines, yet are within the range for bedrock and alluvial streams.
- Because the values of a are not typically reported since this is simply a scaling factor we didn't include that in the text. The values we got for a were 16.8 in the sand channel and 8.5 in the fine grained substrate.