Response to Referee Comments

We would like to formally thank the reviewers for their comments and insight. We have revised our manuscript to take their comments into consideration and feel the manuscript has benefitted greatly from the review process. We address conceptual issues and comments that were in common from both reviewers first, and then address detailed comments from each reviewer.

Both referees sought a more comprehensive review of the experimental design and methodology and results from each experimental run. We have revised the Results and Methodology sections to provide additional details regarding our experimental design (e.g., infiltration capacity measurements and process for simulating uplift), analysis (e.g., method used for channel type classification, calculation of basin integration rates, and calculation of channel incision rates), and results of individual runs versus aggregated observations based on experimental conditions. We also include an additional figure depicting the channel type classification methodology (Fig. 4) and a digital supplement with imagery for all runs like Figure 6 in the current manuscript.

Both referees noted confusion regarding the delineation and definitions of contributing versus non-contributing areas (NCA). Referee #2 highlighted that NCA neglects to consider the fact that groundwater divides can extend far into topographically-flat areas. Our use of the term "non-contributing area" only refers to surface waters, and we clarify that in the first paragraph of the introduction. We completely agree that the groundwater flow to channels can deviate from topographically-defined contributing areas, especially in low-gradient environments, and that areas labeled as NCA at the surface can contribute groundwater to channel heads. In fact, groundwater contributions from areas delineated as non-contributing area were likely critical in driving channel development. The presence of groundwater-driven processes in our experiments underscores the potential deficiencies of viewing NCA in the landscape as hydrologically isolated from observations of the surface topography alone. We have revised the manuscript to try and make it more clear that CA delineated through surface topography only refers to surface contributions and that subsurface divides are not constrained by surface divides.

Referee #2 felt that the manuscript was biased towards overland flow processes for two reasons: 1) contributing areas (CA) and non-contributing areas (NCA) are delineated by surface topography alone, and 2) sediments were saturated prior to the start of the experiments. As noted above, NCA was specifically defined to focus on surface water contributions. One of the main conclusions was that groundwater plays an important role in channel network growth and thus NCA determination as it is traditionally done does not account for this important contribution. We make that more clear now. We do want to be clear that defining CA and NCA based on surface water contributions alone does not bias the experiments towards overland flow processes. The definition of sub-watersheds eroding predominantly via overland flow processes was defined based on slope, relief, and morphology of channel heads, and then compared with upstream CA and NCA. By keeping the definitions of CA and NCA focused on surface water only, it helps us better distinguish the potential drivers of overland flow (i.e. greater CA). Unfortunately, we do have a way to distinguish the subsurface drainage divides and thus cannot compare directly the subsurface contributing area to channel heads that might be connected specifically to seepage erosion.

We also want to address Referee #2's concerns regarding saturating the sediment before initiating our experiments noting that it may bias the results towards more overland flow. This is a really good point, and we acknowledge that it likely did bias the results towards overland flow at the start of each run. Later in the runs, once more relief had been established, the partitioning of flow between overland flow and subsurface flow as mediated by infiltration capacity and substrate erodibility is more informative, and we still feel that the results later in the runs are not biased by initial conditions.

The decision to saturate the sediment was intentional. Three conditions made it necessary to saturate the sediment before beginning the experiments: (1) infiltration rates exceeded the rainfall rate under all conditions, (2) the initial topography was flat, and (3) no uplift had been applied to the system (i.e., base level drop via lowering the outlet's gate). An unsaturated experiment would likely exhibit an initial lag time with little to no channelization since surface flow processes would be inhibited by the flat topography and high infiltration. Given no initial uplift, seepage erosion would also be precluded because the outlet would lack a seepage face for groundwater to discharge from and entrain sediment. Therefore, to facilitate the immediate formation of channels, we chose to saturate the sediment from the experiment's onset. The referee is correct that initial saturation does bias the experiments toward overland flow at the beginning of the experiments since some amount of initial uplift could have been applied to provide a seepage face. The manuscript has been updated to note this, both in the methods section and in section 5.1 of the discussion.

Several comments concerned scaling relationships between our experiments and the real-world. The model we used was designed to be a process model rather than a scale model, to demonstrate whether varying conditions could lead to different processes of channel development. Previous experiments have focused predominantly on providing ideal conditions for either surface water or groundwater-driven processes of channel development. Our experiments sought to provide a middle-ground with suitable conditions for either process to occur, uniquely demonstrating how both processes could co-evolve in the same low-gradient drainage basin. Although morphologies may be disproportionate to those found in the real-world, we feel the conclusions derived from inferring processes from morphologic indicators, observing the integration of NCA through time, and delineating surface water drainage areas using topography remain valid. Revisions to the manuscript include a new paragraph at the start of the discussion on scaling.

Referee #1 – General Comments (*in italics*)

"The captions of the figures are not complete. Ideally, a figure should be understandable (at least the main point) just by reading the caption. Often the reader does not [know] which run is tested on each figure."

> Response: The revised manuscript includes revised figure captions to provide greater context for interpreting and understanding the importance of data presented in each figure.

"The method to determine the contributing and non-contributing area is not clear for me. A graphical explanation may be useful."

> Response: The manuscript includes Figure 3 that demonstrates the process of delineating contributing and non-contributing area. The text preceding this figure (lines 233-247) was revised to better couple with the language and imagery shown in the figure.

"*I do not understand also, what is the quantitative criterion to attribute a channel to overland flow or seepage. The numeration should be used consistently*."

> Response: The current manuscript includes Table 3 listing the quantitative criteria for attributing a channel to overland flow. The criteria were consistently applied across all runs. The revised manuscript includes an additional figure showing how these criteria were applied to classify channels (Fig. 5) and revised text in lines 265-275. In addition, we made it more clear where the

cut-off thresholds are for determining if a channel head is mapped as overland flow (Type 1) or seepage (Type 2).

"*Moreover, some of the conclusions are not really supported by the data. Which curve does demonstrate each conclusion point? Some behavior occurs for only one run, so it may be incidental. As the number of runs is small, all should be shown with the same kind of plot found in Fig.6, maybe in appendix*."

Response: The revised manuscript has been updated to relate our conclusions with an associated figure. The authors agree that the chance of incidental behavior is greater due to the small number of runs. Therefore, the revised manuscript now includes an appendix with imagery like (former) Figure 6 to provide readers with the opportunity to independently evaluate each run. In addition, throughout the discussion, we have tried to make it more clear which data support each statement made.

"*The run 1 is closer to the previous experiments of the literature. Do the authors find the amphitheater-headed channels in this run? I would say, that the elevation of the watertable is significantly smaller than the total bed, which produces mass wasting under the form of slumping or sapping events. At least one example of mass wasting should be shown and illustrated*."

Response: The referee is correct that run 1 is the most similar to previous seepage erosion experiments as it was conducted with relatively cohesionless sediment. Yes, amphitheater-shaped channel heads attributed to seepage erosion were observed in run 1, as depicted in Figure 8. Additional imagery included in Fig. 4 which shows both sapping and overland flow channel heads were present. A digital supplement was added that includes DEMs from all runs to further demonstrate examples of mass wasting events associated with seepage erosion channels.

"*I do not understand also the discussion of the field examples. In Fig. 13 and 14, in which case seepage erosion is dominant (likely right and left). The morphological indicators must be indicated in the caption. I note, also there is no discussion about the scaling between laboratory experiments and the field. Can we deduce the relevant time and space scales using the laboratory results? Are the shapes similar?*"

Response: The imagery in Figures 13 and 14 has been updated to highlight the key morphological features demonstrated by the field examples. Likewise, the caption was updated with a brief discussion of the morphological indicators. The scaling issue is addressed up above and covered in the text in lines 478-492.

Referee #1 – Specific Comments

1.) "*The uplift process is not sufficiently described and discussed. Can the authors show in the schema of Fig. 2, how this uplift is applied? Consequently, does the main slope evolve with time*."

Response: There are three issues here. The first involves how uplift is accomplished. Uplift of the basin is accomplished by lowering base level. Because that could cause confusion, we have edited the text to make it more clear that we are lowering base level at the outlet. The second issue is how that is being done. This is described in depth on lines 189-194. We have included more detail in the caption for Figure 2. In addition, Figure 2 is revised to show the gate mechanism and associated step-motor used to control base level at the outlet of the system more clearly. The caption and supporting text have been revised to explain how lowering base level at the outlet using the gate mechanism effectively applies uplift to the system. The last issue is whether the main slope evolved with time. The main channel slopes are a function of the amount of base level fall and the distance of the channel head from the outlet. It is possible this was constant over time, but the main channel slope was not evaluated as part of our analysis.

2.) "*Is Table 3, obtained for a specific run? If not, are the channels head similar for all parameter values? A graphical example would be worthwhile, to understand the procedure*."

Response: The values presented in Table 3 were extracted from several runs during early, middle, and late timesteps to capture a representative range of values for both channel types. The channels selected for criteria extraction were judged to be the most characteristic examples of seepage erosion or overland flow channels that could then be applied to more ambiguous channels. The revised manuscript includes an additional figure (Fig. 5) demonstrating this procedure and the text in the paragraph on lines 265-276 was revised to be more clear.

3) *How the NCA integration rate is defined and then computed from experimental data? Same question for the incision rate?*

Response: The NCA integration rate is defined as the area of NCA converted to CA per hour. The rate was computed by differencing the area classified as CA in the evaluated timestep from the area classified as NCA in the preceding timestep. The resulting area measurement was then divided by the total time between scans to derive a rate. This text has been added into the methods section following the description of how NCA and CA are defined.

The incision rate is defined as the depth of sediment eroded per hour. Differencing the elevation of sequential timesteps provided the depth component of eroded sediment (Figure 12) which occurred over a known length of time between scans, providing a rate. Computing incision rates for overland flow and seepage erosion channels required aggregating an average incision rate for each channel type per timestep. This clarification is now added to the methods section in the paragraph describing volumetric erosion rate calculations (lines 297-305).

Referee #2 – General Comments

"*1. The paper is biased towards overland flow type processes. The definition of Non Contributing Areas (NCA) does not take into account the fact that groundwater divides can extend far into topographically flat areas. Thus much of the discussion on hydrologic connection applies to surface-water connection only and, although this lowers the ambition of the paper, it should be clearly acknowledged or challenged*."

Response: The comment was addressed in more detail up above. We were only able to quantify surface water connections as we only evaluated contributing and non-contributing area as defined by the topography. The authors agree that groundwater divides can extend far beyond surface water divides, especially in topographically flat areas. The implications of this regarding drainage network evolution are discussed in lines 59-69 of the Introduction.

"*2. Experimental conditions and runs have to be more precisely described especially the way infiltration rates were measured.*"

Response: The revised manuscript includes an expanded Methodology section describing how we measured infiltration rates for different sediment compositions. This updated description is on lines 208-212.

"We measured infiltration capacity of each sediment composition using a single ring infiltrometer constructed of a 30 cm long cylindrical tube. The tube was placed vertically over a bed of pea gravel to allow for drainage and loaded with sediment to a thickness of 15 cm. After saturating the sediment, water was then added to the tube to a depth (head) of 10 cm. The time needed for the falling head to completely infiltrate the sediment was recorded, allowing the infiltration capacity to be calculated. The test was repeated a several times, and the average value was reported in Table 1."

"*3. The authors discuss the difference between two channel types but we lack a description and characterization of these channels. We should be shown various examples of these channel types, their characteristics should be discussed, and so for the evolution from one type of channel to the other."*

Response: The revised manuscript includes an expanded discussion of channel type characterization. Specifically, a figure with a photograph of a run with both Type 1 and Type 2 channels is included (Fig. 4), and a figure demonstrating the classification process for the two channel types in greater detail is included (Fig. 5). Additionally, a digital supplement with imagery for each run, like Figure 8, has been added to allow readers to independently evaluate each experiment.

"*4. Eventually I would suggest a revised version could be more focused on the experiments, their description and analysis in order to provide a useful account of these runs*."

Response: The revised manuscript provides a more thorough discussion of the experimental methodology and results of each run. Specifically, we include more information on infiltration rates (lines 208-212), how we handled side-wall influence (lines 243-244), how the NCA integration rate was calculated (lines 262-265), how the channel head delineation criteria were set and applied (Lines 268-272 and 274-276), and how incision and volumetric erosion rates were calculated (lines 297-305). In addition, we added in a description of the general behavior of each run in section 4.2 in the results section (line 340-363).

Referee #2 - Specific Comments

(Line 17) "in the abstract you mention that seepage and overland flow channels occur for different infiltration and precipitation rates, then you tell that seepage occurs after overland flow channels had developed. This is a bit confusing"

Response: The revised Abstract clarifies that seepage erosion and overland flow did not occur exclusively under a given set of conditions. Rather, the experiments had co-evolving channels by both processes, which developed to a greater or lesser extent depending on the experimental conditions and amount of relief generated by channel incision over time.

"Seepage-driven erosion was favored in substrates with higher infiltration rates, while overland flow was more dominant in experiments with high precipitation rates, although both processes occurred in all runs."

(Line 20) "What do you mean by surface water contributing area for seepage channels?"

Response: Although groundwater drives seepage erosion, the channels maintain a small surface water contributing area defined by the topography.

(Line 32-35) "The notion of NCAs as internally drained areas seems a little bizarre to me. This view seems oriented by some surface flow vision. I can't see how these NCA can't be connected through groundwater flows to the network that is going to incorporate them. Given the setup the entire system has a single groundwater drainage whose boundaries are the cylinder walls."

Response: As stated in the General Response section, the authors agree that NCAs may not be hydrologically isolated. We specifically define NCAs with respect to surface water contributions, and thus internally-drained means that surface water drains into those areas instead of into the expanding channel network. The water that enters areas defined as NCAs then either infiltrates and flows as subsurface flow into the channel network or flows into the channel network via spillover events as water levels rise. The balance between those two pathways is controlled by both precipitation and infiltration rates. In the natural world, water in NCAs could also be lost through evaporation and transpiration but that is unlikely in our experimental set-up. The impermeable bottom and sides of the tank do direct groundwater to drain at the channels, which increases the likelihood that NCAs are contributing groundwater to channels. These points are clarified in the revised manuscript.

(Line 55-57) "Again this claim only applies to surface flows, at least at the beginning of the experiment."

Response: Agreed, the following paragraph, lines 58-68, presents the case for hydrological connections between NCA and channels via surface or groundwater.

(Line 60) "I would be curious to know how these spillover events have been recorded in natural settings."

Response: Hilgendorf et al. (2019) summarized the literature regarding spillover events and how they have been recorded in natural settings. Generally, these events are recorded by fluvial systems that flow across topographic barriers (transverse drainages). Spillover is a mechanism by which channels could breach topographic divides and form a persistent outlet. A classic version of spillover events occurs frequently in wetland and lake complexes throughout the glaciated Central Lowlands of the northern U.S. Wetlands or lakes that are internally-drained under most conditions spill into channel networks during high precipitation events that drive levels over low divides. In most cases, those connections are transient, but incision from repeat events can make those connections permanent.

(Line 65) "Why should hydraulic conductivities be contrasting?"

Response: The cited studies were conducted in glacial settings with deposits characterized by complex assemblages of heterogeneous sediment. Depositional environments range from gravelly outwash deposits with higher hydraulic conductivities compared to clay-rich glacial tills. Successive glacial advances and retreats deposit layers of sediment with varying compositions, thus contrasting hydraulic conductivities are present at depth.

(Table 1) "How do you measure the infiltration capacity I?"

Response: The manuscript has been updated to include the following description of infiltration capacity measurements. We measured infiltration capacity of each sediment composition using a single ring infiltrometer constructed of a 30 cm long cylindrical tube. The tube was placed vertically over a bed of pea gravel to allow for drainage and loaded with sediment to a thickness of 15 cm. After saturating the sediment, water was then added to the tube to a depth (head) of 10 cm. The time needed for the falling head to completely infiltrate the sediment was recorded, allowing an infiltration capacity measurement to be calculated. The test was repeated a dozen times, and the average value was reported in Table 1.

(Table 2) "The "low rainfall" rate is 8μm/s ∼ 29mm/h which turns out to be a very large rainfall rate. For an experiment of 10 hours this means a rainfall on order of 30cm... This does not often happen in real life even in equatorial settings. The "high rainfall" rate then corresponds to ∼ 60cm for ten hours which is within a factor of two from the world record for a precipitation of that duration."

Response: Agreed, even the lowest rainfall rate used in the experiments were greater than rainfall rates found in the most humid environments. However, such rainfall rates are not uncommon when using physical models. Hasbargen and Paola (2000) and Gazzetti (2015) used a nearly identical apparatus with rainfall rates of 6.5 μm/s and 17-24 μm/s, respectively. Other experiments using different apparatuses had similarly high precipitation rates: Lague et al. (2003) (28 μm/s), Ouchi (2011) (10.5 μm/s), and Turowski et al. (2006) (12.5 – 39 μm/s), to name a few. Rainfall rates are high to provide consistent channel-forming discharges that allow drainage networks to form on the order of hours. Although it would be rare to find a storm system that delivered rainfall intensities of 3 cm/hr for 10 hours, it is not uncommon for rainfall intensities of that magnitude to fall for short durations of time. By keeping rainfall rates high during the entire duration of the experiment, we allow for continuous erosion throughout the experiment.

(Table 1) "It seems as if no erosion and network growth can occur if there is no continuous uplift. Then, how can you decorrelate the importance of this factor compared to the others? what if there was no uplift or at least if $U/R << 1$? Did you test this?"

Response: We did not conduct experiments without uplift or with a lower uplift rate. We expect that a run without uplift would lack channelization given the flat initial topography. The basin would have simply filled with water until spilling over at the outlet with no means of incising channels. Over an equivalent timespan to other experiments, a much lower uplift rate could have resulted in some important differences in channel evolution. Given the importance of channel incision in creating relief throughout the basin and enabling seepage erosion, a lower uplift rate could have generated less relief and inhibited seepage erosion channels from forming. We maintain that since the uplift rate was constant across all runs, we can decorrelate its influence at the given rate.

(Figure 2) "How do you prevent water flow from the tank sides and if you do not prevent it how does it affect the experiment?"

Response: Water flow along the tank sides did occur during the experiments. This was mitigated to some extent by placing barriers along the upper lip of the tank to prevent water pooled there from spilling into the basin. However, there was some water flow that incised channels along the edge of the tank during some runs. When this occurred, those channels and associated drainage areas were excluded from analyses. We now mention this earlier in the methods section (line 223-226).

"Some precipitation collected on the walls of the experimental drum, which could influence channel development along the edges of the experimental basin. To account for this possibility, all drainages along the edge of the basin were removed from digital terrain analyses."

(Line 205) "What would happen if the system was not saturated before the start? Given the Infiltration rate you describe I suspect you would have no overland flow and only sapping processes. Did you try? Did sapping channels form and develop?"

Response: See our General Response addressing concerns about saturating the media.

(Line 225) "The definition of NCA and CA as zones where surface waters do or do not contribute is made clear here. It should probably appear earlier."

Response: Agreed, the revised manuscript notes these definitions earlier, in the abstract and the first paragraph of the introduction.

(Line 235) "Same comment for the influence of flow on the walls. It should appear earlier in the description of the experiment."

Response: Agreed, the revised manuscript notes the influence of water flow along the tank walls in the Methodology section (see above).

(Line 250) "Can you show examples of each type of channel head please so we can make up our minds?"

Response: The revised manuscript includes a new figure (Fig. 4) with an image from Run 3 with both kinds of channel heads present. We also include a digital supplement with time series DEM data for each run. In addition, all data are in Sockness and Gran (2021), in a public data repository as noted in the Data Accessibility section.

(Table 3) "What is the rationale behind these choices and again can you show us the channel types you are talking about?"

Response: The values presented in Table 3 were extracted from several runs during early, middle, and late timesteps to capture a representative range of values for both channel types. The channels selected for criteria extraction were judged to be the most characteristic examples of seepage erosion or overland flow channels that could then be applied to more ambiguous channels. The revised manuscript includes an additional figure demonstrating the channel classification methodology for each channel type (Fig. 5). Time series data for all runs are now in a digital supplement.

(Figure 6) "The Channel type 2 you seem to show here looks more like a mass flow then a sapping channel. Do you have other examples? Are these channel types always localized on the borders?"

Response: Yes, other runs formed channels via seepage erosion. The revised manuscript now includes a digital supplement with time series DEM data for each run depicting other examples. Seepage erosion channels were not localized to the borders exclusively.

(Line 324) "You say that type 1 overland flow channels comprise the vast majority of channels during the first half of the experiment but you force this to be so by saturating your media before the start of the experiment. This does not seem to be natural. Then could you show some example pictures of the changes from type 1 to type 2 you describe? I would really be curious to see how run 3 looks like in the end."

Response: See our General Response addressing concerns about saturating the media. The revised manuscript includes a digital supplement with time series data for each run that depicts the transition from type 1 to type 2 channels. In addition, Fig. 5 illustrates in more detail how a channel head could shift classification from 1 to 2 during a run.

(Line 338) "[W]hat sort of unit is this m2 · m−2? Is that not just the same as a fraction? Can you explain?"

Response: Correct, it is a fraction. We included units to ground fractional values in their physical basis.

(Figure 8) "There are maybe two issues here that are pervasive in the entire manuscript. First you are not clear about the difference you make between CA aka contributing area and CA aka channelized area (sic). [I]s there any and if so can you be more specific. If, as it seems at first, they are the same then a second problem arises as [we] do not know what is the real basin area of seepage channels. You always stick to your surface flow definition of the contributing area. But this has not much meaning for a channel fed by groundwater. You should be clearer on these issues and potential limitations."

Response: Contributing and channelized areas are two distinct components of the watersheds in our experiments. Contributing area refers to the upland (non-channelized) area that contributes surface water to channels. Channelized area refers to the area occupied by a channel: It does not include the area upstream of the channel heads that drains into the channelized area. The abbreviation for contributing area, CA, defined in line 224, is used only when referring to contributing area. The revised manuscript: (1) distinguishes between contributing area versus channelized area more clearly and (2) more clearly states that contributing area defined by surface topography does not necessarily represent the groundwater contributing area.

(Line 420) "and discussion thereafter [referee emphasis] Again because your initial setup favors a specific type of process (overland flow channels) it seems difficult to be able to sort out the influence of the factors you describe with your initial saturation. It really seems important that you discuss this point somewhere."

Response: See our General Response addressing concerns about saturating the media. The revised manuscript includes a discussion of the implications of saturating the media before initiating the experiment. The main issue is that it may lead to a bias in overland flow at the start of the experiments. Later in the experiments, when there is enough relief for seepage erosion to start occurring, the initial saturation of the basin likely did not have an impact on erosional rates. We have clarified that in the discussion now.

(Line 455) "Again if one looks at table 1 the infiltration rates you give should prevent infiltration because they are much higher then your rainfall rates (which are already huge). Thus this discussion seems a bit far-fetched."

Response: We assume that "prevent infiltration" is an error, and the referee intended to state "prevent overland flow/runoff/surface flow." We disagree that infiltration rates exceeding rainfall rates preclude overland flow, which is described by the process of saturation overland flow (i.e., saturation from below). Given saturated sediment, the water table is near the surface throughout the basin and is nearly coincident with the channel surface. Constant rainfall across the basin raises the water table as it infiltrates, causing the water table to reach the surface, especially in near-channel areas. Additional water input to these saturated areas travels as overland flow. Increasing or decreasing the clay content scales the degree to which this occurs in addition to modifying sediment cohesion.

(Line 465-470) "I'd love to be persuaded but clearly you do not disclose the evidences needed to claim this."

Response: The conclusions regarding rainfall rates, clay content, and erosion volumes are supported by empirical measurements. The conclusions regarding groundwater flow patterns are more interpretive because we had no means of measuring groundwater flow in our experiments. However, the presence of active seepage erosion coupled with small surface contributing areas support the role of groundwater flow in driving channel network development even though we could not measure the subsurface contributing areas directly.

(Line 499) "Again I think that because you saturate your media prior to the start your experiment is biased towards overland flow and therefore conclusions about clay content and rainfall are probably far-fetched."

Response: See our General Response addressing concerns about saturating the media.