

Response to Reviewer #3:

First, we would like to thank reviewer #3 for taking the time to review our submission. Reviewer #3 provided general, as well as line-by-line comments, which we will address below. While the original review comments are shown in *italics*, our responses are given in regular blue font.

First, I enjoyed reading this well written manuscript. I appreciate that the authors crafted an accessible background literature review (from the perspective of a nonexperimentalist). In their manuscript, Tofelde et al., develop interesting and timely scientific questions and knowledge gaps—what are the responses of alluvial fill terraces to modulation of base level, and changes in upstream water discharge and sediment supply (Q_w , Q_s respectively)—which they then address using seven experiments. I echo the sentiment of Reviewer 2 that this paper has the ring of a review paper, yet that is not a problem for me, and I actually appreciated the good explanations of current knowledge (theoretical, field, and experimental). I thought the amount of review in the introduction was appropriate to bring a non experimentalist/expert up to speed on the current thinking of how terrace incision-aggradation functions with respect to changes in upstream or downstream (base level) boundary conditions. I thought the figures are well made and that the captions are effective as well.

Thanks for this kind assessment.

The results of the seven experiments performed by the authors show there are distinct responses in the slope of pre-perturbation and post-perturbation alluvial surface elevations that are dependent upon the type of forcing mechanism, and the authors document interesting transient behavior of fill terrace, channel elevations/width, and Q_s out of the experimental system with time. In experiments with increased Q_w or Q_s , gradients in the new equilibrium channels decrease significantly compared to the pre upstream perturbation channel gradients. This is a somewhat intuitive, yet interesting result, and one that presumably has the potential to be tested in the sedimentary/geomorphic record. I thought that the rationale for the experiments and the results are thought provoking to those interested in not only morphologic response of alluvial fill terraces to external forcing, but also the implications of their response to external forcing in terms of chemical signatures preserved (or not) in sediment/sedimentary systems (end of Section 5).

The experimental design did not include simulations of increased $Q_w + Q_s$, or decreased $Q_w + Q_s$, as conceivably might occur/be expected in a natural sedimentary system undergoing upstream changes in boundary conditions. Thus its possible the C2 results of these experiments (pure perturbations in Q_w or Q_s) may be difficult to invert from sedimentary records or be more pronounced in experiments than nature. I don't consider this a shortcoming of the manuscript, it's just an observation, and perhaps the authors could include a statement about this in the discussion?

We agree that changes in environmental conditions (e.g. tectonics, climate) that have the potential to affect either Q_s or Q_w are likely to affect both in reality. For example, a change to wetter conditions (increase in Q_w) might also trigger a pulse of sediment release from the hillslopes to the channels (e.g. Steffen et. al (2009, 2010)). Thus, considering the entire sediment routing system, Q_s and Q_w are often coupled. With our experimental set-up, however, we only investigate the response of the transfer sub-system to changes in surrounding conditions, and we de-couple Q_s and Q_w to investigate the potential effect that each of those two parameters can have on the evolution of channel morphology. Also, although both parameters are thought to vary simultaneously, thick fluvial fills and fill terrace formation in the field are often related to either significant changes in either Q_s or Q_w (hillslope-driven and discharge-driven models as described in Scherler et al. (2015); see p.3 l.27 to p.4 l.4). As such, we investigate the

two end-members of those models. Many variations in-between those endmembers are possible though. For clarification, we will include the above mentioned points within the manuscript.

Other reviewers have suggested ideas to help improve the communication of what results are novel by the restructuring of the literature review and parts of the discussion (e.g. Malatesta's comment #2). I concur that the authors should consider improving the way in which they communicate how to interpret these experimental results in the context of existing theoretical and experimental knowledge.

We agree with both reviewers that the introduction on the theoretical background should be extended. Please see our reply to Malatesta's comments for details on how we intend to adjust the section on background knowledge.

Recommendation: I recommend that this manuscript ultimately be accepted for publication after the authors implement minor revisions.

Line-by-line comments:

P1 L9 suggest "...tectonic histories" rather than "...tectonic conditions"?

We appreciate this suggestion, however, we prefer to stick to the term 'tectonic conditions' for the following reason: Terraces form under certain environmental conditions. As such, the terraces can be used to reconstruct those certain conditions that persisted at a certain point in time. They are not a continuous archive (as for example a varved lake core would be). Therefore, fill terraces cannot be used to infer entire climatic or tectonic histories.

P2 L20-21 You may want to specify that (at least for Schaller et al 2004) the methods used to interpret paleo discharge were in part based on cosmogenic nuclide concentrations, not simply the age of terrace formation. Interpretations from those concentrations are in turn subject to assumptions of the systematics of cosmogenic nuclides and sedimentary dynamics.

The main point about this sentence was to state that fill-terrace deposits have been used in various ways, including for example the reconstruction of paleo-discharge or paleo-denudation rates. Schaller et al. (2004) did not reconstruct discharge, but paleo-denudations rates. If we explained the Schaller work in detail, we would also need to explain the other applied approaches, which would not benefit the purpose of the sentence. As such, we prefer to leave the sentence as it is.

P3 L10 The following sentence needs to be rewritten: "To our knowledge, there are no experimental studies that systematically compare how fill terraces formed through various mechanisms may differ from one another, or investigate the impacts of terrace formation on downstream sediment discharge."

We hope that the following adjustment of the sentence will help to clarify its structure: "To our knowledge, there are no experimental studies that systematically (i) compare differences in fill terrace geometry, location, and formation timescales that occur for various formation mechanisms, or (ii) investigate the impacts of terrace formation on downstream sediment discharge."

P5 L33 The end of the second Section (2 Formation of fluvial fill terraces) seems abrupt; would it help to provide one or two statements that help summarize and transition into Section 3 here?

For a better transition, we will add the following sentence at the end of section 2: "In summary, different processes within the sediment-routing system, including variability in sediment supply, water discharge, base-level changes, autogenic processes, or complex channel responses to a single allogenic perturbation, could lead to similar geomorphic responses of the alluvial river - the formation of fluvial fill terraces."

P8 L2 Suggest “channel incision” rather than “river incision”?

OK, will be changed.

P9 L23-26 “When comparing terrace slopes to the active channel slopes (blue lines) at the end of each run, terrace slopes are steeper in all experiments in which upstream conditions (Q_w , Q_s , in) were changed 25 (Fig. 6 A-D). In contrast, the slopes of the terraces and the active channel in the BLF experiment are similar to each other (Fig. 6E).” This is a really interesting relationship, and one I would not have expected (though I don’t often think about these kinds of experiments), but that does seem intuitive. Is this pre-perturbation terrace slope and upstream-downstream boundary condition relationship something that is seen in other experimental studies? In nature? I see your discussion includes some mention of this explicitly, and introduces the active tectonic aspect that unfortunately complicates interpretations and adds non uniqueness to potential interpretations of terrace slope history. Can you predict/offer guidelines for which kind of natural systems your experimental results would be best applied?

Variability in terrace slopes has been reported from field studies (e.g., Tofelde et al. (2017), Baker and Gosse (2009), Burgette et al. (2017), Poisson and Avouac (2004)). As such, if (1) base-level changes, autogenic processes and complex responses can be ruled out as the driving formation mechanism of those terraces and if (2) the terrace surfaces have not been tectonically deformed, then those fill terrace slopes can potentially be used to infer paleo Q_s and Q_w . The idea of paleo-hydrology inferred from terrace surfaces has already proposed by Leopold and Miller (1954). With modern techniques, however, we hope to improve the possibilities for more quantitative assessments of Q_w and Q_s of the present and the past. We will include this potential application of the findings within the discussion.

P10 L22 add a space after “...Fig 5).”

OK.

P15 L30-31 Perhaps cite the figure # again for clarity, for which grey vs. yellow circles relate to this sentence.

Will be done.

P16 L6-8 The last sentence of Section 5 suggests chemical signals may be propagated more efficiently through systems during phases of aggradation, rather than phases of incision when mixing of older stored sediment might overprint the chemical signature of “fresh” hillslope derived sediment. This is interesting... Your statement makes sense, however would it also be fair to say that the chemical signature would be a function of the ratio of the “fresh” to recycled sediment (and obviously the erosion rate upstream)? And that those ratios could vary greatly given different system scales (I’m thinking about the ratio of upstream derived Q_s vs excavated volume)? Perhaps this is a tangential idea more suitable for its own paper?!

We appreciate this suggestion to consider absolute ratios between upstream supplied and remobilized sediment. As we have the absolute input and output Q_s values, we can generate such a plot, which we will include in the lowest panel of Fig. 5. We will adapt the text of the manuscript accordingly and include the point that the degree of signal modification is dependent on the ratio of fresh vs. recycled sediments.