

I have now carefully reviewed the responses of the other reviewers and the responses of the authors. I commend the other reviewers for their attention to the detailed history of the literature on grain-size transitions, while noting that my main focus was closer to the idea of bank control on the transition between gravel-bed and sand-bed rivers, which is what the authors have now clarified as their intent. I believe that the relative lack of criticism from my part was due to my realizing this focus. In hindsight, it may also be a direct result of having been investigating the same problem myself, and having come to the same general conclusion about the influence of the cohesive strength of muds on channel pattern, just a few weeks before I received this invitation to review. This primed my brain for the paper, and indeed contributed substantially to my enthusiasm for the work of the authors, but meant that my focus on the problem at hand caused me to not notice some of exposition required to set up the general background. Therefore, I am glad to have seen the more critical reviews of the other referees as well as the response of the authors and their revised manuscript.

I remain in agreement with the conclusion that around the sand–gravel transition, the width of rivers with sandy sediment is controlled by the cohesion of bank materials (this force being stronger than the weight of the particles), and those with gravel beds is controlled by the weight of the particles. I find this to be a simple and mechanistic line of reasoning with the ability to explain and unify a number of observations – many researchers have danced around this point, but this is the first time that I have seen it explicitly stated. I think that the work is worth publishing on the merits of this alone.

However, I do not agree with the authors' decision to forge ahead with their characterization of all gravel-bed rivers as threshold-state rivers while not noting that above-threshold gravel-bed rivers are a major conclusion of Millar and Quick (1998), and in fact refusing to cite the recent demonstration of this by Pfeiffer et al. (2017). In this, the authors' views are at odds with the evidence. I had thought that above-threshold gravel-bed rivers were simply omitted by accident the first time around, which lead to my earlier and more positive review; this omission could be easily addressed by citing the studies appropriately and noting that *most* gravel-bed rivers are near-threshold rivers. What I have trouble understanding is that the main addition of the submitted paper to the scientific knowledge, and the one that captured my enthusiasm (and continues to do so) is the gravel–sand transition. This makes this gravel-bed river threshold-state issue somewhat beside the point at hand, and one that unfortunately will cause me to withdraw my support for publication of this (in my opinion) otherwise-sound paper until it is addressed.

Comments from the prior review that were not addressed by the authors:

Use `\citep{}` instead of `\citet{}` with extra parentheses.

Please do this. Or maybe you were just using `\cite{}`? Whichever way it is, please just use BibTeX properly.

Comments from prior reviews that were not satisfactorily addressed by the authors:

p.3, l.22-24. Note also that Pfeiffer et al. (2017) demonstrate that gravel-bed rivers are significantly above the threshold of motion in rapidly-uplifting settings...

The authors state that they are unable to reproduce these results and therefore will not cite the paper. I do not find their arguments to be convincing for the following reasons:

1. The authors claim an issue with the statistical treatment used by Pfeiffer et al. without demonstrating what this is or how it might change the results (see response to AE Turowski on C2 and C3). Claiming an error on a piece of existing literature requires evidence.
2. The authors claim that their conclusion is based on the biases involved in a slope-dependent τ_c^* ; I have tested their data set with both a constant and slope-variable (Lamb et al., 2008) τ_c^* , and have obtained a range of τ_b^*/τ_c^* from approximately 0.3 to 10 in both cases.
3. The authors argue that there was no independent measurement of τ_c^* in those rivers with high τ_c^* : I have not checked this but τ_c^* typically changes over a factor of 2 whereas τ_b^*/τ_c^* (at bankfull) ranges from ≈ 0.3 to ≈ 10 . This therefore cannot explain the observations. Furthermore, to be fair to Pfeiffer et al. (2017), it is important to note that τ_c^* is often not measured independently in the field, making it difficult to do so with such a large compilation, and that for rivers in which τ_b^*/τ_c^* is great, such observations could be difficult with available technology at best and dangerous at worst.
4. The authors incorrectly state that the authors do not provide data on hydraulic geometry. The supplementary information has full data on hydraulic geometry. The authors likewise state that the authors do not give information on discharge; while this is true, the data provided (slope, depth, width, grain size) can, with an appropriate roughness formulation, can be used to generate a reasonable estimate of discharge.
5. Finally, the authors claim that they were unable to reproduce the results of the Pfeiffer et al. study and failed to do so. However, it is unclear based on this response how the authors did this and/or what data sets they used (Pfeiffer's? Their own?) and whether the data set used for validation indeed was sufficient to test the ideas put forward in that paper. They claim that this means that the results cannot be replicated under "more rigorous testing". From my point of view, Pfeiffer et al. is a peer-reviewed paper with data that support the conclusion that gravel-bed rivers exist at above-threshold transport stage; the claimed "more rigorous testing" on the other hand, has no description to support it. The answer clearly must be to stick with Pfeiffer et al. until/unless future research provides an alternative explanation for their findings.

While I appreciate that a thorough analysis of the Pfeiffer et al. (2017) study may go beyond the scope of the paper, it seems that one is left with two choices. Either the authors should at least tacitly accept it and the idea that rivers may not be as tidy as they would like, or they must thoroughly disprove this. The former would require just a sentence noting that they are not looking at environments of rapid uplift, and therefore are not evaluating the explanation put

forward by Pfeiffer and co-authors. The latter would require a more extensive explanation. In my opinion, it is better to look at the publication of the two papers by Phillips and Jerolmack (2016) and Pfeiffer et al. (2017) as a challenge of how to look deeper into the data and validate whether or not the discrepancy exists, how to improve our measurement abilities, how to critically evaluate our assumptions, and (if necessary) how to unify theory.

p.11,l.31-32: Would you like to discuss some of the reasons for the low frequency of channels with 1-10 mm grain size? In particular, do you think that this may have to do with the crystal size / granule break-down problem, or possibly be connected to the transition between cohesion-dominated banks and particle-weight dominated banks that makes these grains either difficult to move or whisked away in a larger-clast gravel-bed river? This is of course ignoring arguments for equal mobility...

The authors state that this point is not important to this paper. However, there remains a fundamental question: does the bimodal transport state exist as the result of a bimodal input, a bimodal filter within the system, or both? The authors seem to argue for the middle answer, but do not mention that it may instead be, for example, the result of a bimodal input to which the internal response of the system that the discuss herein is more a response and less a driver.

In the authors' response to Reviewer 3 on the use of a constant Chézy friction factor, I have some open questions.

First, how did they calculate the friction factor in the response? This is not stated.

Second, friction will change with the presence of bedforms. The authors explicitly do not address form drag. This is OK by me since they state this clearly. However, I question at this point whether a constant Chézy friction factor selection (0.1) is to look for deviations from a particular relationship, or if it is, as the authors state, that they simply do not care about reducing scatter. Could you please clarify this?

Line-by-line comments:

p. 1, line 15: (I did not know this during the first review round – sorry to bring it up only in round 2) Lacey was actually the first to generate a power-law relationship for hydraulic geometry. A nice review is by Savenije (2003): “The width of a bankfull channel; Lacey’s formula explained”.

p. 1, lines 20-21: I am not going to hold you to this because you are trying to generalize (as is appropriate here), but I will note that there is significant scatter in the power-law hydraulic geometry relationships, and would suggest that the scales of the scatter vs. the strengths of the relationships are appropriately acknowledged.

p. 2, line 5: in a rectangular channel (again, I missed this the first time around, but is easy to note)

p. 8, lines 27-29: Phillips and Jerolmack put significant effort into measuring τ_c^* and compiling measurements of it along gravel-bed rivers in the mostly-

tectonically-inactive mountain west of the USA and in Puerto Rico. (I know they included a river in the Oregon Coast Range, but study the low-relief portion of the river.) Their study excludes data from rapidly-uplifting landscapes required to address the Pfeiffer et al. (2017) argument as well as some of the references brought forth by Reviewer 3 on non-threshold behavior. Therefore, the authors have described only part of the sum of the knowledge, and have put forward a statement about threshold behavior based only on this. The cautionary note here is that it is easy to say “most rivers do this” but hard to say that “all rivers do this”. The difference between “most” and “all” could hold some important information into the forcings and response. I appreciate that these may be beyond the scope of this paper, but I find it important to avoid such blanket statements and partial referencing of the literature on a particular problem that can result in a disjointed scientific literature.

p. 12, lines 1-2: The Kean & Smith reference does not support the statement about mud, vegetation, and erosion thresholds. (I missed this as well in round 1 because I had not yet read this paper, and therefore did not realize that it was mis-cited.)

p. 12, lines 32-33: Your paper is a really nice piece of work about the channel width transition between sand- and gravel-bed rivers. You do not address the question at all about all alluvial rivers being near-threshold (see above comments).

p. 14, lines 1-3: Such a statement may be true, and has been rephrased to be consistent with Millar and Quick’s work. However, without an evaluation of the Pfeiffer et al. work, it is not possible to say that that this is true in general.

p. 14, line 3: Contrary to my last comment, this statement is still at odds with Millar and Quick (although their range of τ_b^*/τ_c^* is less than that reported by Pfeiffer et al.). I have gone back and checked this following Lamb et al. (2008), and my calculated τ_b^*/τ_c^* values range from 0.6 to 2.6.

Conclusions (in general): it is possible that some material in the second half of your conclusions could go in your discussion.