It would be helpful to see the learning and washout time series for the 2 subjects that they tested in the constant condition and the 3 in the high variability condition. The preliminary data that are shown only give the mean and std during of SAI during learning. This does not illustrate the time course of learning and the time course of washout trials, the latter of which is their main outcome measure based on model predictions.

Reviewer 1:

R#2

This is a well written stage 1 registered report that proposes a design to test whether and how movement variability (here variability in step asymmetry) affects a form of use dependent locomotor learning. The paper largely relies on a behavioral paradigm that was described in a recent article by Wood et al. 2020 and two different computational models. The two models are shown to respond differently to increased variability-one is a use dependent model with a strategic component added on, and the other is a Bayesian model.  
  
I have no major concerns about the hypothesis being tested-- it is interesting and timely. However, there are a few things that would be worth thinking through a bit more, or justifying a bit better, within the design.  
  
1) The Wood et al. 2020 paradigm used a gradual introduction of a perturbation during learning and the experimental design proposed here uses an abrupt change during learning. These can result in different after-effects in other types of motor learning paradigms for walking (e.g. adaptation). Are there reasons to think that these types of perturbations would be equivalent in this use-dependent experiment? This may influence the power analysis since it is being done using gradual learning data and applied to abrupt conditions. It might not be a problem, but it seems worth thinking about.

Simulate this, address this and given thought, maybe we can add another condition?   
  
2) I also assume that you switched to an abrupt change so that the model fitting would include both adaptation and de-adaptation. You state that you will model individual subject data, which is appropriate, but I am not clear why you will model all three conditions combined? Wouldn't it be a stronger test if you model one condition for each subject (e.g. the high variability condition where you expect the greatest differences) and then see how those model parameters apply to the other conditions? Can you clarify?

We wanted to ‘stack’ conditions to get a single set of parameters for each subject. This reviewer is suggesting that we pick a condition and model it, and see how well it predicts the other conditions  
  
3) The use-dependent + strategic model seems to be based on the use-dependent model from the Diedrichsen et al. 2010 paper. Correct me if I am wrong, but it adds in an assumption about a strategic component that was not in the Diedrichsen paper-namely that there is a retention factor for the strategy that is assumed to be used from one stride to the next. I would like some more intuitive justification for the need for the strategy component and for fitting the A parameter. In walking, subjects may easily have time to modify the gait pattern online and hit a target, thus they might only need the C\*en part of that equation. It would be nice to understand the basis for the A\*sn component. It is not entirely intuitive. Perhaps it just biases the model in the direction of the abrupt perturbation? More explanation would be useful. Perhaps even a plot showing how the different components of the model change as a function of stride, which might help the reader intuit.

Could make the empirical argument that UE studies use . Also provide a conceptual argument, . could also be thought of as a proxy for your confidence or the stability in the environment. Fine with adding plot.  
  
4) The simulations that you show have a high SD for the learning phase in the groups where variability was added. It makes me wonder what the individual fits might end up looking like? The fit relies so heavily on the learning portion of the data since it is nearly half the data for each condition. Do you have individual subject examples? Perhaps I am missing something?

This should be easy to provide arguments for this.