**Review Radtke et al. PCI RR**

Thank you for giving me the opportunity to review this Stage 1 registered report proposal. To start off this review, I would like to point out that I am not an expert on the details of processing and analyzing pupillometry data, so I hope that one of the other reviewers was able to review that aspect of the proposal. The authors attempt to clarify the role boredom plays in effortful self-control. This is a very interesting question given that many cognitive tasks we use to study effort and self-control are highly repetitive and thus it is plausible that boredom might be a confounder. The thing I am not convinced by (yet) is whether this study will be able to disentangle effort and boredom, partially because I did not fully understand what the authors consider effort and boredom to be. I list my comments for the authors to consider below.

1. One thing that does not become clear to me is how the authors think conceptually about effort and boredom. On p. 3, the authors talk about effort as both a behavior (investing effort or working hard) and a subjective experience (feeling of effort during mental activity). Later on, we see that they plan to measure effort both via self-report and pupillometry. I am guessing that these two operationalizations fit onto the two different efforts (self-report for subjective experience and pupillometry for investment of effort) but this is not explicitly explained or justified. How does the investment of effort relate to the experience of effort, and how do both relate to boredom? On p. 5, the authors talk about boredom mainly as a subjective experience (being bored). Much of what they write about the experience of boredom on p.5 also applies to the experience of effort (reflect the cost of ongoing actions and to signal that maybe we should do something else). Thus, while effort and boredom feel different and often occur under slightly different circumstances (mostly people experience effort when task difficulty is high and experience boredom when task difficulty is low, though there are exceptions as the authors note), they are though to have a similar purpose. The authors then propose that experiencing boredom can be effortful. This can make sense when it comes to how much effort needs to be invested, though I think we would need to think hard about how to gain confidence in the idea that it is the perception of boredom that leads to the higher effort investment and not something else that co-occurs with the rise of boredom. I don’t think it makes sense to conceptualize that experiencing boredom can be effortful in the subjective sense as these are both experiences that are thought to track a similar thing, so I would not know how to disentangle them. In summary, I think it is important that the authors clarify how they think conceptually about effort and boredom and are clear in what ways they might relate to each other. I believe this paper is a great example for clearly distinguishing the investment of effort from the feeling of effort: <https://doi.org/10.1016/j.concog.2018.05.013>
2. Thought probes: Related to my first point, I am concerned about the self-reports regarding effort due to task difficulty and effort due to boredom. The authors need to clarify how effort could result from boredom. Also, it is important to ensure that these items have validity. Do people have insight into effort that arises from task difficulty and effort that arises from boredom? Do we know how participants answer these items and that they insight into this process so that they can self-report on it? Have previous studies used these items and validated them in some form?
3. Analyses: Higher perceived task difficulty in HCT, higher perceived boredom in LCT, lower performance in HCT: these are all clear and basic effects to expect. Thought probes: As I mention above, I am not sure what boredom-related effort and task difficulty-related effort are and whether people are able to report on these processes, so I am not sure what this analysis can tell us. Pupil size: Again, I think the authors need to present some evidence that we can have confidence in the self-reports of boredom-related and task-difficulty related effort. Also, the authors cannot rely on explained variance to compare the models. Adding more variables to a model will always increase the explained variance, but it also increases the likelihood of overfitting the data. For that reason, model comparison needs to penalize complexity in some form. This could be done by comparing models with an information criterion (AIC/BIC/WAIC) or via cross-validation.
4. Interpretation given different outcomes: This makes sense but rests on the assumptions that a) we can disentangle boredom-related and task difficulty-related effort and b) participants are able to report on this. These assumptions are crucial and I believe need to be explored to ensure that this very interesting project will succeed in telling us something new about boredom during effortful self-control.

I hope my review is helpful and good luck moving forward.

Jonas Dora, University of Washington