We would like to thank the reviewers for the encouraging and valuable comments on our paper. We structured our response into six categories:

1. Method criticism (R1, R3): *R1 criticised the lack of a learning algorithm, and R3 questioned the practical utility of our approach.*

The key goal of our study is to clarify the objectives that underlie neural activities in higher-order brain areas (such as the prefrontal cortex). We do so by showing that one objective—efficient compression of task-spaces—gives rise to the type of activities measured in these areas. The questions of how to learn those representations, or how they could be utilized, are, of course, of high interest. However, they really constitute a second or even third step. Before asking them, we need to first establish a clear, mathematical objective, i.e., what exactly prefrontal areas are trying to achieve. It may be surprising, but currently no such objective exists.

2. Conceptual novelty (R3): *R3 questioned the conceptual novelty of our submission.*

Our main conceptual contributions are three-fold: First, we provide a clear hypothesis for the function of higher-order brain areas, which we support with a comparison to neural activity. Second, we consider both state space ('model-based') and policy ('habitual') compression. Third, we link the type of compression (state-space vs policy) to the animal's behavior. We thank R3 for the references (which we will include), but note that both Poupart et al, 2003 and Bertsekas, 1995 only consider the state-space-compression case, not the policy compression case. Indeed, policy compression is a novel concept to our knowledge, and the two compression strategies have not previously been compared.

3. Habitual vs model-based systems (R1): *R1 remarks that the habitual system requires a model-based system, which is deemed wasteful.*

R1 raised a very deep question: why have a separate habitual system, given that one will need a model-based system anyways. While we do not have an answer to this question, we note that this is exactly what brains seem to do: they have both habitual and model-based systems that are thought to either cooperate or compete in order to enable efficient control (for a recent discussion see e.g. Kool, Cushman & Gershman, 2018).

4. Match to data (R1, R5): *R1 finds that the model of the somatosensory working memory task does not match the data because it misses ramping dynamics. R5 asked for a more quantitative comparison of data and model.*

The mismatch of data and model noted by R1 must be a misunderstanding. The lower panel of Fig 5B, for example, shows a model neuron with ramping up activity towards F2. Moreover, the second, condition-independent demixed PC (Fig. 5D) also shows ramping during the delay period. Indeed, to really compare data and model, one has to compare the population data, which we here do using demixed PCA. (We would like to emphasize that we mainly show the single neuron examples because this is the standard when assessing models of these tasks, see e.g. refs [14] and [17]). We completely agree with R5, though, that one would eventually want to have a clear metric to compare model and data. However, at this point in time, most models of higher-order areas such as the PFC fail to even demonstrate a qualitative match to data. Second, and maybe for that reason (!), there is currently virtually no established protocol to compare neural data to model data for these higher-order areas. We note that these areas pose their own problems because of the flexibility and dynamics of their responses, so that establishing a metric is a non-trivial task. In the absence of such a protocol, we therefore prefer to simply visualize the match.

5. Additional analysis: *(a) R1 requested analysis of reaction times in the delayed licking task. (b) R5 suggested to model a model-based behaviour task. (c) R3 requested a comparison of the compressed and history state. (d) R5 suggested to model the delayed licking task with a linear system.*

(a) This is an excellent suggestion, but unfortunately, the recorded data for the delayed licking task does not allow to analyze reaction times (Inagaki et al, personal communication.) (b) Also a great suggestion, but unfortunately, there is currently no well-established model-based behavior for which neural activities have been recorded. A key problem here is that it is very hard to elucidate whether an animal is behaving habitual or model-based (see e.g. Akam et al., 2015). (c) In Figures 3,4 and 5, we have a qualitative comparison of the compressed state with the history state. Fig 5G, e.g., shows how the compressed state lies in a subspace of the history state. (d) Linear systems are too rigid to capture this task, because one action may influence the next action through the state, and because actions are in general a non-linear function of state. This is an example why in the future it will be important to look for more powerful parametrizations.

6. Technical problems: *R3 finds that eq. 4 is trivially true.*

In eq.4, the two distributions that are marginalized over are different: one is factored according to the history policy and the other according to the compressed policy. We understand R3's confusion and will make this dependency more explicit.