We thank the reviewer for carefully reading our revised manuscript and for providing useful suggestions on how to improve it.

We have expanded our description of t-SNE as suggested. t-SNE only preserves local distances but not global distances. Thus, while it is a useful tool to cluster objects by similarity, the distances between clusters in the t-SNE projection are not necessarily meaningful because global distances are not preserved. We now mention this both in the main text and in the caption. We also explain in the main text that we use t-SNE primarily as a visualization method. To more rigorously quantify how well representations of odor identity are preserved, we calculate the mutual information (MI) between novel foreground odors and ORN responses in the 50-dimensional space. We also provide a short explanation of what MI is in the main text.

We added explicit description of how the stimulus is estimated in the main text together with details about the definition of the signal perturbation. We also state clearly in the main text that the result is an estimate of the magnitude of each signal component and therefore both the identity and the intensity of the signal are estimated.

We have revised the text at the location indicated in the review to address the reviewer's concerns. Note that we hold the adaptation timescale at 250 ms, but vary the memory timescale (longer ones are more helpful). We also added a phrase in the following paragraph to provide intuition as to why longer time scales of memory are useful.

We have now added on Figure 4c, the results for the non-adaptive case, showing that in absence of front-end adaptation, primacy sets are not well maintained, particularly when the primacy order is small.

This is an interesting observation. We have rewritten this part of the text as suggested and have mentioned in the main text the interesting observation made by the reviewer. We have also expanded the comparison of our results with those of Olsen et al.

We rephrased this to mention other types of adaptation, but removed the explicit implication of the role of phosphorylation sites.

mechanistically, it seems that the minimum background at which adaptation begins is set by spontaneous channel opening. If that is correct, it is not hard to envision models in which the onset of adaptation is independent of beta. Could you be more specific in this paragraph about the extent to which the conclusion about small deviations from WF behavior being beneficial depends on specifics of the model?

We added a phrase to emphasize that this effect is more general than our specific model.

my understanding has always been that combinatorics (i.e. a single odorant molecule activating multiple receptors) helps resolve the problem of dimensionality. Hence I felt that the setup for this section was confusing because it raised an issue that I thought was resolved, or at least less of an issue than implied here.

Combinatorics does imply that many odors can be encoded within ORN responses, but does not necessarily state how the odorants could be decoded from the ORN responses. CS gives a framework in which combinatorics and sparsity allow one to decode both odor identity (which odorants comprise odor) and individual odorant intensities. We edited the first sentence of this section to make this clearer.

there is some repetition in the paragraph starting on this line (e.g. there are two, slightly different, versions of the cost function).

We simplified the text.