nature research

Peer Review File

Manuscript Title: Evolving schema representations in orbitofrontal ensembles during learning

Reviewer Comments & Author Rebuttals

Reviewer Reports on the Initial Version:

Referees' comments:

Referee #1 (Remarks to the Author):

Although orbitofrontal cortex (OFC) is known to make an important contribution to reward-based learning, the exact nature of this contribution remains unclear. There are considerable challenges to studying this issue. The behavioral problem being learned must be sufficiently complex to engage OFC. This typically requires training across many days, but it is difficult to record the activity of the same neurons across multiple days, particularly given the position of the OFC deep in the brain. In the current study, the authors solve these problems with an elegant behavioral paradigm and a sophisticated data analysis that enables them to examine how OFC constructs a schema of a behavioral task across multiple days, which it can then use to speed learning on subsequent problems. The ability to generalize is a central component of intelligent behavior and so these results will be of interest to a broad range of scientists, including neuroscientists, cognitive scientists, and researchers developing artificial intelligence.

A problem with recording a neuronal population across multiple days is that one cannot guarantee that the same neurons are being recorded. To overcome this, the authors use dimensionality reduction in order to capture the most relevant dimensions of the neural response, and then canonical correlation analysis to ensure that neuronal responses recorded on different days are aligned so as to maximize the correlation between dimensions across days. Not only is this a neat algorithmic trick, but it is also of theoretical significance, suggesting that OFC activity as a whole is evolving towards a common representational space, such that it is irrelevant exactly which neurons are recorded from day to day. Despite this being a relatively complex analysis, it is presented very clearly with helpful diagrams and simulations. Indeed, the entire paper is very well written. Consequently, I have largely minor comments.

Main comment

My main comment relates to the results shown in Figure 3f. I think these are really interesting as they hint at the representation space favored by OFC. However, they receive minimal discussion, and I'm not sure that I follow the authors' very brief explanation. They state that OFC is favoring the representation of "hidden task variables", by which they are referring to the odor sequence and the consequences that has for correct performance of the task. However, all positions involve hidden task variables, because there is nothing intrinsic to the odors that indicate whether it will be associated with reward or not. The animal has to learn this information. So there must be more than OFC simply favoring hidden variables. Maybe I'm missing something in my understanding of what the authors mean by a hidden variable. Either way, it would be nice for this finding to receive a little more attention in the manuscript.

Minor comments

1. Page 7, last sentence of the first paragraph: there is an extraneous "that".

2. Fig. 2a and 2b: I'm assuming these are arbitrary units? The color axis ought to be labeled, including whether they are indeed arbitrary units.

3. Fig. 2d: I don't think this is very clear. I can see an increase in explained variance of the first dimension with training, but I wouldn't be able to read-off this information for the second or third dimension. I'd suggest doing a pseudocolor plot, with explained variance as the color axis.

4. Fig. 3c and 4c: The color axis needs labeling as "Decoding (%)".

5. Fig. 5a and ED7a: I believe the r-value is the correlation between the grey and black lines. This should be stated in the legend.

Referee #2 (Remarks to the Author):

Zhou et al. use a previously published odor sequence task to study schema formation in the OFC. The most challenging element of the task are Positions 4 and 5 in Sequence 2, where the odor valence is dependent on the previously sampled odors. Nonetheless, rats are able to learn this task and variants on the task where the task structure is the same, but the odor identities are new (Problems). The ability to use a common task structure with different stimuli is a clever method for studying schemas. Next, the authors use a variety of computational analyses to derive dimensionality and representational similarities of OFC activity during the task. Over the course of learning, dimensionality decreases and population activity becomes more similar for similar elements of the odor sequence. They also use a state-of-the-art manifold alignment analysis to infer low-dimensional activity across days even when recording from separate neuronal populations. They find that the Problems are encoded more similarly late in training compared to early, suggesting that the OFC representations grew less idiosyncratic and more general, consistent with schema formation. The authors additionally show that this analysis yields similar results across rats, suggesting that the OFC utilizes a generalized neural code. Finally, they show that behavioral learning of subsequent Problems and OFC representations accelerate with more experience. The topic of study is interesting and timely. The analytical techniques used are innovative. However, there are concerns related to statistical analyses, interpretation of the results and novelty.

Major Concerns
1. Statistical rigor. A large portion of analyses were poorly explained and many statistical tests were missing where necessary. For example, Fig 1d-g all need statistical tests. Furthermore, several figures were missing error bars or it was not clear why error bars were absent (Fig 2e-g, Fig 3e-f, Fig 4e-f, Fig 5b-d). In addition, figures were missing labels (Fig 2a&b, Fig 3c- what does the colored-bar scale represent, Fig 3b- y-axis, more below under Minor Comments).

2. There is a general lack of clarity about the data content in numerous figures (e.g., Fig 5, which is a separate bullet point below). The odor sequence task is extremely high-dimensional (odor identity, valence, Position, Problem, rat). Particularly since individual data points were rarely plotted it was frequently unclear what data was actually being run through each analysis. For example, what data was used in Fig 2? Data from one rat? Data from multiple rats and one Problem? All rats and all Problems? There is very little information about sample size (n=?). Being explicit about the inputs to the analysis is particularly important in order to understand what is actually being shown. For example, if the data in Fig 2b spans multiple Problems, that could mask effects of learning and the authors should pursue separate Mahalanobis distance measurements separately for each Problem and perhaps even for each rat. Similar criticisms apply to the majority of the figures.

3. Schema or reward learning? The authors find that OFC representations become generalized with training and interpret this to mean the development of schema. Another interpretation is that the OFC learns reward contingencies. Other papers have shown sensitivity to reward in OFC (Farovik

et al., 2015). The authors themselves have published on OFC's role in assessing value in their 2019 Current Biology paper. Admittedly, it is difficult to dissociate reward learning and schema formation in rodent studies. In their favor, a case could even be made that there is no distinction between learning reward contingencies and the formation of a schema. To a rodent, their most important schemas may revolve entirely around rewards. Can the authors make a case that schema formation is or should be considered a distinct function from reward learning? To be clear, the value of these results does not depend on an affirmative answer. Regardless of their answer, the results are potentially interesting and novel, but whether "schema" applies here should be carefully considered.

4. The authors say that "the OFC is required for cognitive mapping" or schema learning, however, there is no direct manipulation of the OFC to show it is necessary for schema learning in this task. All evidence was based on observations of neural activity as it relates to behavior, which is correlative in nature. To say that OFC is required, the authors would need to test for necessity with some gain of loss experiment. Otherwise, the claim needs to be adjusted to better fit the evidence.

5. Figure 5 is confusing. The authors cleverly use poke latency under different "expected reward" conditions as a way to measure the rats' ability to compound on prior memories of the task structure. However, there were a number of elements to this analysis that were either not explained or too briefly explained. First, why are only the first 5 trials plotted in Fig 5a? Is the rest of the analysis only using the first 5 trials (again, see bullet point #2)? Why is Next+1 included as a condition? Is there evidence that the rats are looking two trials ahead? This seems unlikely. For Day 1 or Day 15, is (-, -, -) significantly different from (-, -, +)? The text describing dimensionality compression should briefly explain how it was calculated or at least a summary of what it is. I have no idea what "Pattern evolution" is and why it relates to cross-rat decoding in any way. Finally, the analyses in Fig 5e seem unorthodox (correlating areas of the curve of various metrics across time), particularly since the timescales of Fig 5b and Fig 5c,d are different. Significantly more attention should be paid to this figure since it has the unique job of demonstrating application of a schema to multiple Problems.

6. There are a couple of papers that were published related to the present work that were not cited. Morrissey et al., eLife, 2017 from Takehara-Nishiuchi's group showed that over weeks following the acquisition of two distinct associative memories, neuron firing in the rat prelimbic prefrontal cortex became less selective for perceptual features unique to each association and, with an apparently different time-course, became more selective for common relational features, implicating medial prefrontal cortex in forming a schema. The other relevant paper is Rubin et al., Nature Comm, 2019 from Yaniv Ziv's group. They showed that prefrontal cortex revealed schematic representations of distances and actions, and more importantly, that the internal structure was conserved across mice, allowing using one animal's data to decode another animal's behavior.

Minor Concerns
The peri-event time window should be included in the main text. Related to the peri-event time window, why are there 8 time points when there are 6 epochs to a trial (light, poke, odor, unpoke, choice, and outcome)?

The authors could indicate that there were (at least) 15 days in between Problems since the description of the training schedule in the main text is a little vague: "After shaping on the apparatus with an initial odor problem, the rats were trained on five new problems."

The notation for Sequences should remain consistent. For example, the authors use "S1a" in Fig 1b but "1a vs 1b" in Fig 3f.

Figure 1d should highlight which positions contain odors that conflict depending on the Sequence (P4 and P5 in S2).

In Figure 1g, please comment on why accuracy on S2b4- is consistently worse than S2a5-.

The text describing Figures 1h-j needs more explanation. It's unclear what the authors are trying to convey.

Figure 2b should indicate which pixels correspond to positions with odors that conflict across Sequences.

Figure 2a has interesting structure for the first two odors such that the population distinguishes between Sequences (high Mahalanobis distance between S2a and S2b, because they are different odors, presumably). However, this effect disappears on Day 15 because the OFC learns that odor identity is meaningless information in the short-term (the rat is immediately rewarded regardless of odor identity). However, odor identity is meaningful in the long-term (it determines whether the rat should lick on P4 or P5 for S2, yet the OFC loses discriminatory power on Day 15. To me, this suggests that the OFC is primarily concerned with valence of the stimulus (consistent with the point raised in Major Concern #3). Otherwise, wouldn't it be expected that Mahalanobis distance be high for S2a1/S2b1 (in Matlab indices, [3,4]) and S2a2/S2b2 ([7,8]) to reflect the rat's understanding that the odors presented in those Positions are informative? Please comment. Related, In Fig 3c Day 15, the CCA-aligned decoding confusion matrix also seems to confuse S2a4 with S2b5 and S2b4 with S2a5, again consistent with the valence interpretation of OFC representations.

In Figure 3b, the axis labels are not informative. How are the trial types ordered? Which ones correspond to which Sequence and Position? The authors should also indicate which trial types corresponding to S2a4, S2a5, S2b4, and S2b5. Also, if I understand this correctly, the y-axes are not CC #s. Those should be the title of each subplot. What then is being plotted on the y-axis?

Why does the misaligned condition still have prominently high decoding values along the diagonal?

In Figure 3f, why does P6 also become more discriminable over time?

An inherent weakness of using extracellular tetrode recordings is the inability to hold cells across days. How many cells were "double counted" across days and how might that affect the CCA analysis?

Top of page 7 should read "while decoding at positions that were discriminable". Also in that paragraph, "representing hidden task variables that in the first few days of training" has an extra "that".

Referee #3 (Remarks to the Author):

Summary

The authors present results from a study in which rats learned a complex task that required responding or withholding a response based on a sequence of presented odors. Rats learned four sequences; in two, decisions could be made purely based on the currently presented order, while in the other two sequences decisions had to take into account previous items in the sequence. After rats were well-trained on this task, they were presented with a set of 5 new problems with different odors while unit activity was recorded from neurons in lateral orbitofrontal cortex (OFC). As presumably different units were recorded on different days, the authors aligned activity for different problems by reducing the dimensionality of the population data and then projecting the data into a manifold that maximized correlation between two projected data sets. This procedure

performed both for pairs of problems, to align data collected for a given animal, and for groups of rats. Within this aligned data, the authors found that during learning the population data in OFC came to distinguish between individual odors less. Instead, OFC developed a low-dimensional representation that distinguished between trials with different reward values. The dimensionality of the representations was greater for the more difficult sequences that required taking prior history into account. Trial type could be decoded across problems, suggesting that OFC formed similar representations of the task across problems. Similarly, trial type could be decoded across different rats after alignment, suggesting that different animals have similar representations in OFC.

Evaluation

There is substantial interest in neural coding of "schema" representations. This manuscript applies recently developed methods in a novel way to provide insights into how generalized representations are learned and represented in the brain. The study therefore addresses a question of broad interest. However, there are some limitations in the present manuscript that should be addressed. One central issue is that it is not precisely clear what information is contained within the OFC representations quantified in this manuscript. The manuscript would be improved if the authors could provide a more precise definition of schema and how the task structure is represented within the OFC schema. To what extent we should be surprised to find that cross-problem decoding works in this case, given the existence of task-sensitive activity in OFC for each individual problem? As no specific models of how the task schema might be represented are tested, it is hard to know how inevitable it is that different problems are represented similarly. For example, does OFC simply represent the reward on each individual trial, or does it also represent position and more abstract distinctions between diverging sequences? Is cross-problem similarity equally present for different aspects of the task (e.g., position, reward, sequence)? The manuscript focuses mainly on the reward dimension of the task structure. Without greater information about how OFC representations other task dimensions, it is not clear if OFC is truly representing a "schema" comprising multiple task dimensions.

Major issues:

1. While evidence is presented that OFC represents a task schema that is common to different problems, it is less clear what information is represented in the schema that might drive performance. The results mainly focus on the distinction between rewarded and non-rewarded trials. What about other task-relevant information? For example, is the sequence (S1a, S1b, S2a, S2b) decodable? If so, at which positions? Does decoding generalize across positions? Does OFC represent the hidden structure in this task, or could it adequately be explained as representing only the reward at individual trials? There also appears to be position coding to some extent; characterization of this coding would help clarify not only the content of the schema but the role of OFC in representing different types of task content. Prior work has examined similar questions, but they are also relevant here. Particularly relevant is whether the different generalizable aspects of the task (i.e., features other than odor) are all transferred between problems, or whether only some aspects are represented in a generalizable way.

2. The correlation between the neural measures and poke latency on each problem is unconvincing. There are only five observations and the variability between problems is almost entirely captured by two groups. The data would be better summarized as follows: compression, evolution, and behavior all increased after problem 1. Moreover, comparing the changes by trial within each problem would be more useful for testing for a relationship between brain and behavior.

3. The general approach used for the individual analysis techniques is generally well explained, but the details used for specific results is sometimes unclear. For example, I can guess at what the "pattern evolution" matrix represents (presumably a correlation between trial type dissimilarity matrices on each day and the day 15 matrix), but it is unclear from the text. In a few places, I was

unclear what "data iterations" referred to.

4. How did the manifold alignment between problem pairs allow for generalization to different problem pairs? If I understand the analysis, the manifold alignment seems very dependent on the ordering of the canonical correlations. The first three components seem well matched, which makes sense as the same dimensions may explain more variance for all problems, but how many other components were included in the classification analysis? How correlated were each of the component pairs? Analysis of the main components in terms of their sensitivity to task features and their correlation across problem pairs may help to address point 1.

Other issues:

Figure 1j: Markers should be shown for all lines or none of them, and markers should be smaller if included.

Extended Data Figure 1: Reaction time should be split into correct and incorrect trials for the reward trials, or just correct trials included if there are not sufficient incorrect trials to estimate reliable statistics. For the non-reward trials, only incorrect trials should be shown. Currently, statistics for the non-reward trials are a mix of correct (2 s) trials and incorrect trials (varying time), which makes this figure hard to distinguish from the accuracy data shown in Figure 1.

Why were the transition probabilities not equalized when transitioning between sequences? What determined those probabilities?

How did performance vary during days 15-23? Was the best day substantially better than the worst day?

The use of cross-ISI to select independent component runs should be explained further.

Page 15: More detail should be given on the "template matching algorithm."

Page 15: Why were the electrodes advanced between problems to obtain different units?

Page 17: What type of multidimensional scaling was used?

I recommend representing the problems with colors that are distinguishable by colorblind individuals.

Author Rebuttals to Initial Comments:

Note: Author rebuttals in blue

Referee #1 (Remarks to the Author):

Although orbitofrontal cortex (OFC) is known to make an important contribution to reward-based learning, the exact nature of this contribution remains unclear. There are considerable challenges to studying this issue. The behavioral problem being learned must be sufficiently complex to engage OFC. This typically requires training across many days, but it is difficult to record the activity of the same neurons across multiple days, particularly given the position of the OFC deep in the brain. In the current study, the authors solve these problems with an elegant behavioral paradigm and a sophisticated data analysis that enables them to examine how OFC constructs a schema of a behavioral task across multiple days, which it can then use to speed learning on subsequent problems. The ability to generalize is a central component of intelligent behavior and so these results will be of interest to a broad range of scientists, including neuroscientists, cognitive scientists, and researchers developing artificial intelligence.

A problem with recording a neuronal population across multiple days is that one cannot guarantee that the same neurons are being recorded. To overcome this, the authors use dimensionality reduction in order to capture the most relevant dimensions of the neural response, and then canonical correlation analysis to ensure that neuronal responses recorded on different days are aligned so as to maximize the correlation between dimensions across days. Not only is this a neat algorithmic trick, but it is also of theoretical significance, suggesting that OFC activity as a whole is evolving towards a common representational space, such that it is irrelevant exactly which neurons are recorded from day to day. Despite this being a relatively complex analysis, it is presented very clearly with helpful diagrams and simulations. Indeed, the entire paper is very well written. Consequently, I have largely minor comments.

Main comment

My main comment relates to the results shown in Figure 3f. I think these are really interesting as they hint at the representation space favored by OFC. However, they receive minimal discussion, and I'm not sure that I follow the authors' very brief explanation. They state that OFC is favoring the representation of "hidden task variables", by which they are referring to the odor sequence and the consequences that has for correct performance of the task. However, all positions involve hidden task variables, because there is nothing intrinsic to the odors that indicate whether it will be associated with reward or not. The animal has to learn this information. So there must be more than OFC simply favoring hidden variables. Maybe I'm missing something in my understanding of what the authors mean by a hidden variable. Either way, it would be nice for this finding to receive a little more attention in the manuscript.

Thank you for reading and commenting on our paper. Your summary is spot on and exactly what we hoped to convey. With regard to our use of the term "hidden" variables, we used this term to try to convey the tendency of OFC to converge on representing the underlying reward-related structure in the task in a way that is orthogonal to whether external events make that structure readily discriminable. This tendency is evident if you look overall at what the OFC seems to encode or represent in well-trained rats. It is not simply representing reward or value, which would compress all of the positions into rewarded or not rewarded, nor is it simply using external information to assign causes (either the current odors or even the odor sequences), since if it were, we would end up with 24 nice unique states or trial types encoded in OFC. Instead the OFC compresses externally different trial types when possible (at positions P3-P6 on S1 and to a lesser extent also for P1 and P2), while also splitting trial types that are very similar when necessary for performance (P3-5 on S2). While this might be simply seen as value learning, we can extract and discard the variance related to value and still decode this structural information, so we think a more parsimonious explanation is that the OFC is, in essence, identifying latent or underlying causes for why reward appears on some but not other trial types, pruning away unnecessary external information here, adding new internal information there. The representations in OFC are all hidden in a sense; it is just that sometimes they do covary with distinct sensory information.

Most of this was shown in a prior paper (Zhou et al, Current Biology, 2019). Here we simply used the knowledge of this pattern as a tool to explore whether its development would obey predictions for a schema – a generalizable form of knowledge about a problem or circumstance – the most critical of which was to show that the underlying neural manifold was converging on a common structure across problems. To show this, we tested whether activity on one exemplar could be used increasingly well to interpret activity on another exemplar. Finding this seems to us to require that whatever is being extracted about the task in OFC is capable of generalizing to new problems; that is, the OFC is representing a general structure. By definition, such a structure cannot be tied to the idiosyncratic cues used in a given problem.

So this is all to explain our meaning with regard to the term hidden variables. We agree it is perhaps imprecise, but it captures to the best extent possible what we think defines what OFC is extracting about the task.

That said, we also agree that it would be an important improvement if we could put labels on some of the dimensions defining the schema-related neural activity space. To this end, we have examined the relationship between a number of task features and the CCs identified in the training and test sets in our cross-problem and cross-subject decoding analyses. This analyses found strong correlations between CCs 1, 2 and 3 and task features related to current trial value, location with respect to the unique or overlapping odors on the arms of the sequences, and position on the sequence.

Unsurprisingly, value is a strong aspect of what is represented, but value-orthogonal aspects of the sequence are also a prominent component of the schema. These new analyses are presented in **Figs. 3b-d** and **4b-d** in the revision and also in **Extended Data Figs. 9, 10, 12,** and **13.** We hope the above explanation coupled with this new information on what the schema actually consists of will satisfy this concern.

Minor comments

1. Page 7, last sentence of the first paragraph: there is an extraneous “that”.

Thank you for pointing this out; we have removed the extra word.

2. Fig. 2a and 2b: I’m assuming these are arbitrary units? The color axis ought to be labeled, including whether they are indeed arbitrary units.

Yes, they are arbitrary units; we have added “a.u.” on the top of color axes.

3. Fig. 2d: I don’t think this is very clear. I can see an increase in explained variance of the first dimension with training, but I wouldn’t be able to read-off this information for the second or third dimension. I’d suggest doing a pseudocolor plot, with explained variance as the color axis.

We tried using a pseudocolor plot with a single color scale, but it did not show the data well. So instead we use multiple color scales to better show the data in the **Extended Data Fig. 8**, while in **Fig. 2d**, we plotted data only for Day 1 and Day 15, two most representative days, for clearer visualization.

4. Fig. 3c and 4c: The color axis needs labeling as “Decoding (%)”.

Thank you for noting this; it has been corrected.

5. Fig. 5a and ED7a: I believe the r-value is the correlation between the grey and black lines. This should be stated in the legend.

Thank you for this suggestion; we have noted this in the legend.

Referee #2 (Remarks to the Author):

Zhou et al. use a previously published odor sequence task to study schema formation in the OFC. The most challenging element of the task are Positions 4 and 5 in Sequence 2, where the odor valence is dependent on the previously sampled odors. Nonetheless, rats are able to learn this task and variants on the task where the task structure is the same, but the odor identities are new (Problems). The ability to use a common task structure with different stimuli is a clever method for studying schemas. Next, the authors use a variety of computational analyses to derive dimensionality and representational similarities of OFC activity during the task. Over the course of learning, dimensionality decreases and population activity becomes more similar for similar elements of the odor sequence. They also use a state-of-the-art manifold alignment analysis to infer low-dimensional activity across days even when recording from separate neuronal populations. They find that the Problems are
encoded more similarly late in training compared to early, suggesting that the OFC representations grew less idiosyncratic and more general, consistent with schema formation. The authors additionally show that this analysis yields similar results across rats, suggesting that the OFC utilizes a generalized neural code. Finally, they show that behavioral learning of subsequent Problems and OFC representations accelerate with more experience. The topic of study is interesting and timely. The analytical techniques used are innovative. However, there are concerns related to statistical analyses, interpretation of the results and novelty.

Major Concerns
1. Statistical rigor. A large portion of analyses were poorly explained, and many statistical tests were missing where necessary. For example, Fig 1d-g all need statistical tests. Furthermore, several figures were missing error bars, or it was not clear why error bars were absent (Fig 2e-g, Fig 3e-f, Fig

4e-f, Fig 5b-d). In addition, figures were missing labels (Fig 2a&b, Fig 3c- what does the colored-bar scale represent, Fig 3b- y-axis, more below under Minor Comments).

We apologize for not providing enough statistical support. In response, we have increased our statistical descriptions in the figure captions in the main text and made available complete statistical results for all of the comparisons in the main text in an Extended Data table (**Extended Data Table 3**). Statistical tests in the extended data figures were added in the captions under each figure.

Labels for color-bar scales and necessary error bars have also been added in the revision for all relevant figures. With regard to the specific questions above, the heatmaps in **Figs. 3e** and **4e** (now changed to **Figs. 3g** and **4g**) represent correlations between confusion matrices from different pairs of days, so there are no error bars where those data are converted into line plots for Day 15. But we have indicated the significance of the correlation with thick markers on the curves. P values were corrected with BH-FDR.

2. There is a general lack of clarity about the data content in numerous figures (e.g., Fig 5, which is a separate bullet point below). The odor sequence task is extremely high-dimensional (odor identity, valence, Position, Problem, rat). Particularly since individual data points were rarely plotted it was frequently unclear what data was actually being run through each analysis. For example, what data was used in Fig 2? Data from one rat? Data from multiple rats and one Problem? All rats and all Problems? There is very little information about sample size (n=?). Being explicit about the inputs to the analysis is particularly important in order to understand what is actually being shown. For example, if the data in Fig 2b spans multiple Problems, that could mask effects of learning and the authors should pursue separate Mahalanobis distance measurements separately for each Problem and perhaps even for each rat. Similar criticisms apply to the majority of the figures.

We apologize for the lack of detail. In response, we have included detailed information in the statistics table (**Extended Data Table 3**) for each analysis indicating what data were used. In addition, we have tried to make more detailed information, such as that cited as missing above, more explicit in the text and figure captions in the revision.

With regard to the specific issues raised here:

In **Fig. 2**, neural data from the five problems and nine rats were aligned to training days (Day 1 – Day 15). We constructed pseudo-ensembles by combining all rats and all problems on each training day to increase statistical power. The analysis with pseudo-ensembles were repeated 500 times. For each repeat, a pseudo-ensemble was constructed by shuffling the trial order within each trial type for each neuron.

On average, ~1123 neurons were recorded on each day. The number of neurons on each day are also shown in **Extended Data Fig. 6**. The number of neurons recorded on each day on each problem and in each rat were shown in **Extended Data Tables 1** and **2**.

Yes, the **Fig. 2b** spans multiple problems as stated above. We also did the same analysis on each problem as shown in **Fig. 5** and indeed we have found a learning effect of dimensionality reduction.

3. Schema or reward learning? The authors find that OFC representations become generalized with training and interpret this to mean the development of schema. Another interpretation is that the OFC learns reward contingencies. Other papers have shown sensitivity to reward in OFC (Farovik et al., 2015). The authors themselves have published on OFC's role in assessing value in their 2019 Current Biology paper. Admittedly, it is difficult to dissociate reward learning and schema formation in rodent studies. In their favor, a case could even be made that there is no distinction between learning reward contingencies and the formation of a schema. To a rodent, their most important schemas may revolve entirely around rewards. Can the authors make a case that schema formation is or should be considered a distinct function from reward learning? To be clear, the value of these results does not depend on an affirmative answer. Regardless of their answer, the results are potentially interesting and novel, but whether "schema" applies here should be carefully considered.

We have two responses to this concern or set of questions.

First, we agree that reward value is an important part of the schema in OFC, however we think it does not explain all aspects of what is represented in the OFC in this or indeed other tasks. This point was made most clearly perhaps in our previous paper (Zhou et al, Current Biology, 2019), in which we dissociated reward value and task structure representations in activity on a single problem in well-

trained rats. Specifically we were able to extract, and discard variance related to value and the remaining variance was still capable of decoding information about the sequences. However in the current paper, there is also evidence that the schema in OFC – as extracted by the cross-problem and cross-subject decoding - is not just representing reward value. For instance, if reward value were the only information extracted by OFC for the schema, the cross-problem/subject decoding in the confusion matrices in **Figs. 3e** and **4e** should show a checkerboard pattern, reflecting the misclassification of all rewarded trial types and all non-rewarded trial types. This is not the pattern observed. Instead positional information is clearly present, even when comparing positions with similar reward values (**Extended Data Figs. 11** and **14**). In the revised manuscript, we now go further by comparing CCs (canonical components; neural dimensions that represent generalized neural activity between problems) extracted by our analyses with various task features, including both reward and aspects related to the sequences (**Figs. 3b-d** and **4b-d**; **Extended Data Figs. 9**, **10**, **12** and **13**). Consistent with our prior paper, we find that reward is well-correlated with the first CC, however the second and third CCs are better correlated with aspects of the sequence such as position or whether the odors were unique or overlapping. This new analysis essentially puts labels on some of the neural dimensions that define the neural space containing the schema. So we hope it directly addresses the above concern.

Second, however, we think that even had we only found value to be represented, we would still argue that the current data is novel and, in fact, goes beyond showing that OFC represents reward contingencies or associations - at least as we would understand them or as they have been shown in prior work. For instance, in Farovik or indeed in all prior studies by our lab, OFC has certainly been shown to represent cue-reward or even context-reward associations. Typically, however, that activity seems to be at least somewhat specific for the cue-reward combinations used – representing conjunctions of cues (or sets of cues) and outcomes. However here we are showing that in addition to such idiosyncratic information, the “neural manifold” in OFC is also extracting higher-order structure that can generalize to new situations involving different external cues. We think this demonstration is quite novel. For instance, to show this, Farovik (or the other studies mentioned below) would have needed to record in entirely new settings (different contexts, novel mazes, etc), and show that neural activity from the original setting could be used to interpret trial structure and predict reward in this new setting. Further we show that this cross-problem decoding accelerates in concert with accelerated learning on new problems by the subject. This coupling between accelerated behavior and cross-problem decoding is the heart of the paper and it is subtly but importantly novel and different from anything our lab or we think others have done previously in OFC.

4. The authors say that “the OFC is required for cognitive mapping” or schema learning, however, there is no direct manipulation of the OFC to show it is necessary for schema learning in this task. All evidence was based on observations of neural activity as it relates to behavior, which is correlative in nature. To say that OFC is required, the authors would need to test for necessity with some gain of loss experiment. Otherwise, the claim needs to be adjusted to better fit the evidence.

We fully agree with the reviewer that the neural data analysis is correlative in nature, and we did not mean to claim OFC is required for either cognitive mapping or schema learning based on the data we presented in this manuscript. In response to this concern, we have gone through the revision and removed any such claim based on our data, although we do still discuss the significance of our results in light of the ample causal evidence from other studies implicating the OFC in exactly those two functions. We hope this is acceptable.

5. Figure 5 is confusing. The authors cleverly use poke latency under different “expected reward” conditions as a way to measure the rats’ ability to compound on prior memories of the task structure. However, there were a number of elements to this analysis that were either not explained or too briefly explained. First, why are only the first 5 trials plotted in Fig 5a? Is the rest of the analysis only using the first 5 trials (again, see bullet point #2)?

Fig. 5a shows examples of how we did the analysis with the first 5 trials on Day 1. We added more trials (Trials 1 – 5; Trials 6 – 10; Trials 11 – 15; Trials 16 – 20) to the analysis in **Extended Data Fig. 15**. For the rest of analysis (**Fig. 5b**), all trials (20 trials for each trial type on Day 1) were used. The behavioral learning curve for each problem over 20 trials on Day 1 was calculated with a 5-trial moving window (step = 1 trial).

Why is Next+1 included as a condition? Is there evidence that the rats are looking two trials ahead? This seems unlikely. For Day 1 or Day 15, is (-, -, -) significantly different from (-, -, +)?

Yes. The poke latency data shows the rats can look two trials ahead. In **Fig. 1h-j**, we did a linear regression analysis with four predictors (rewards on prior, current, next, next + 1 trials.) on the poke latency and found that the current, next, as well as next + 1 trial rewards, but not the past reward, can significantly affect the poke latency. Poke latency on (-, -, -) is significantly higher than that on (-, -, +) for both Day 1 and Day 15 if all five problems and nine rats were combined (Day 1: $p = 0.004$, $n = 37$ sessions; Day 15: $p = 0.02$, $n = 36$ sessions; two-sided rank sum test). This is consistent with our two published papers (Zhou et al. Curr Biol, 2019a, 2019b).

The text describing dimensionality compression should briefly explain how it was calculated or at least a summary of what it is.

The dimensionality estimation was calculated the same way as we did in **Fig. 2**. The major difference is that, in **Fig. 5c**, we did this analysis for each problem. In **Fig. 2** dimensionality was measure as "% of variance" explained by the first three LCs. To make different problems comparable, in **Fig. 5c**, we used normalized "% of variance" (denoted as Dimensionality Reduction Index) for each problem. We have added these details to the figure captions in both **Fig. 2** and **Fig. 5**.

I have no idea what "Pattern evolution" is and why it relates to cross-rat decoding in any way.

We apologize for this term and have changed "pattern evolution" to "schema evolution" which we hope is a more accurate description. We use this term to refer to the increase in similarity between the confusion matrices on different days resulting from cross-problem or cross-subject decoding analyses. This is plotted in the revision in **Figs. 3g** and **4g** and also in the new **Fig. 5d**. In each case, the content of schema (as revealed by the cross-problem/subject confusion matrices) evolved to become more similar with learning (as quantified by the correlation between confusion matrices on different days). It is not specific to cross-rat decoding; it was also in the original paper in **Fig. 3e**, which is cross-problem, but we simply neglected to label it. We have now corrected this, which hopefully will help.

Finally, the analyses in Fig 5e seem unorthodox (correlating areas of the curve of various metrics across time), particularly since the timescales of Fig 5b and Fig 5c,d are different. Significantly more attention should be paid to this figure since it has the unique job of demonstrating application of a schema to multiple Problems.

Thank you for the comment. Because of this comment and also a suggestion of another reviewer, we have removed the original analyses and replaced them with new ones that are focused on the comparisons between Problem #1 and the rest problems. We hope these are more comprehensible.

6. There are a couple of papers that were published related to the present work that were not cited. Morrissey et al., eLife, 2017 from Takehara-Nishiuchi's group showed that over weeks following the acquisition of two distinct associative memories, neuron firing in the rat prelimbic prefrontal cortex became less selective for perceptual features unique to each association and, with an apparently different time-course, became more selective for common relational features, implicating medial prefrontal cortex in forming a schema. The other relevant paper is Rubin et al., Nature Comm, 2019 from Yaniv Ziv's group. They showed that prefrontal cortex revealed schematic representations of distances and actions, and more importantly, that the internal structure was conserved across mice, allowing using one animal's data to decode another animal's behavior.

We apologize for missing these very important papers. Thank you for bringing them up. We definitely agree they are important and relevant, particularly the terrific work by the Ziv lab, so we now cite and discuss them in the text. In our defense, we still believe that our results are unique. While these prior studies show that neural activity develops to capture the important features of an exemplar problem, they do not show that this ability to capture the important features generalizes – and results in faster learning and encoding – on new problems of a similar form. This is the core finding here. We tried to indicate this in our above responses, and we have tried to clarify it in our revised manuscript.

Minor Concerns

The peri-event time window should be included in the main text. Related to the peri-event time window, why are there 8 time points when there are 6 epochs to a trial (light, poke, odor, unpoke, choice, and outcome)?

There are 8 time points (-0.2 s – 0.6 s; bin = 0.1 s) within each task epoch, and there are 6 epochs within each trial. We have included the peri-event time window in the main text.

The authors could indicate that there were (at least) 15 days in between Problems since the description of the training schedule in the main text is a little vague: “After shaping on the apparatus with an initial odor problem, the rats were trained on five new problems.”

Thank you; we have made this change in the revised text.

The notation for Sequences should remain consistent. For example, the authors use “S1a” in Fig 1b but “1a vs 1b” in Fig 3f.

Thank you; we have made this change in the revised manuscript.

Figure 1d should highlight which positions contain odors that conflict depending on the Sequence (P4 and P5 in S2).

Thank you; we have made this change to **Fig 1d**.

In Figure 1g, please comment on why accuracy on S2b4- is consistently worse than S2a5-.

On S2a5-, the rats can use information (reward or non-reward) from the prior trial to facilitate their behavioral performance, whereas on S2b4-, the rats must remember information (the odor) at least two trials back. In practical terms, this means that if they make a mistake on S2a4- and respond for reward but do not get it, they can then use this non-reward to infer they are on the other sequence and recover to make the correct response on S2a5-.

The text describing Figures 1h-j needs more explanation. It’s unclear what the authors are trying to convey.

The data is to show rats could be able to use the sequence information to make predictions about future outcomes. The analysis is also important for further analysis in **Fig 5a-b**. We have tried to clarify this in the revised text.

Figure 2b should indicate which pixels correspond to positions with odors that conflict across Sequences.

Thank you; we have made this change to **Fig. 2**.

Figure 2a has interesting structure for the first two odors such that the population distinguishes between Sequences (high Mahalanobis distance between S2a and S2b, because they are different odors, presumably). However, this effect disappears on Day 15 because the OFC learns that odor identity is meaningless information in the short-term (the rat is immediately rewarded regardless of odor identity). However, odor identity is meaningful in the long-term (it determines whether the rat should lick on P4 or P5 for S2, yet the OFC loses discriminatory power on Day 15. To me, this suggests that the OFC is primarily concerned with valence of the stimulus (consistent with the point raised in Major Concern #3). Otherwise, wouldn’t it be expected that Mahalanobis distance be high for S2a1/S2b1 (in Matlab indices, [3,4]) and S2a2/S2b2 ([7,8]) to reflect the rat’s understanding that the odors presented in those Positions are informative? Please comment. Related, In Fig 3c Day 15, the CCA-aligned decoding confusion matrix also seems to confuse S2a4 with S2b5 and S2b4 with S2a5, again consistent with the valence interpretation of OFC representations.

The reviewer’s description of the data is accurate. The discrimination of specific odor identities indeed declined during learning, which might be one of the causes for neural dimensionality reduction after learning. This is also reflected in cross-problem and cross-subject decoding of sequences at P1 and P2 (**Extended Data figs. 11a-b and 14a-b**). However, although there is a clear decline in discriminability of odor identities at P1 and P2 after learning, the odors are still decodable within each problem, which is consistent with our previous two publications in which we used the same task (Zhou et al, Curr Biol, 2019a, 2019b). Additionally, while discriminating the odors at the first two positions (P1 and P2) is important for later behavioral performance, the OFC doesn’t have to be the most critical brain region to encode such information. Indeed, in comparing information encoding in hippocampus and OFC in the same odor sequence task, we found that, in well-trained rats, at early positions (P2 and P3), hippocampus showed much better decoding of sequences S2a vs. S2b than OFC; while OFC showed better decoding at later positions (P4 and P5).

As to whether OFC is only concerned about the valence of the stimulus, we think there is strong evidence against this throughout the current data. For starters, if the schema in OFC were only about

the valence of the stimulus on a given trial, then the confusion matrices (**Figs. 3e** and **4e**) resulting from cross-problem or cross-rat decoding should exhibit a checkerboard pattern reflecting miscoding across trial types based solely on reward or non-reward. Obviously, this is nothing like the pattern we find.

Indeed, the four trial types at P4 and P5 in S2 exhibited the best cross-problem mapping among all 24 trial types (**Figs. 3e** and **4e**); this cannot be explained by current value because these positions share the same values with many other trial types. Our previous study (Zhou et al, Curr Biol, 2019a) showed, after the removal of value selectivity, the decoding of these four trial types was still exceptional, suggesting information encoding that is orthogonal to value.

However we appreciate the concern and hope it will be allayed by new analyses in the revision, in which we examined the relationship between the initial CCs and relevant task information (**Figs. 3b-d**, **4b-d** and **Extended Data Figs. 9, 10, 12,** and **13**). While these analyses found that value is clearly important, several other dimensions defining the neural space were well correlated with task features not directly equivalent to value.

In Figure 3b, the axis labels are not informative. How are the trial types ordered? Which ones correspond to which Sequence and Position? The authors should also indicate which trial types corresponding to S2a4, S2a5, S2b4, and S2b5. Also, if I understand this correctly, the y-axes are not CC #s. Those should be the title of each subplot. What then is being plotted on the y-axis?

We apologize for this confusion. The order of trial types in this figure is: P1(S1a,S1b,S2a,S2b), P2(S1a,S1b,S2a,S2b), P3(S1a,S1b,S2a,S2b), P4(S1a,S1b,S2a,S2b), P5(S1a,S1b,S2a,S2b), P6(S1a,S1b,S2a,S2b), laid out in the same order as in all the other plots in the prior figures. We have made this explicit in the figure and captions. The four key trial types (S2a4, S2a5, S2b4, and S2b5) were highlighted. We put the CC#s as the title and labeled y axis as “CCA Score”.

Why does the misaligned condition still have prominently high decoding values along the diagonal?

The random shuffling of trial types still occasionally permits correct or more frequently semi-correct alignment to persist; this causes the above-chance decoding. With this rather conservative approach, the mean decoding accuracy under the aligned condition is still much higher than that under the misaligned condition (**Figs. 3e-f** and **4e-f**).

In Figure 3f, why does P6 also become more discriminable over time?

The decoding of S2a vs. S2b at P6 indeed becomes better during learning, while doing so is not required by the task. We think this happens because OFC encodes past reward history (S2a and S2b have different rewards at P5). Alternatively it may reflect the slight imbalance in the probabilities of transitions between subsequences in our design. The same data is now presented in **Extended Data Fig. 11b**.

An inherent weakness of using extracellular tetrode recordings is the inability to hold cells across days. How many cells were “double counted” across days and how might that affect the CCA analysis?

We did CCA analyses across problems or across subjects (but not across adjacent days). It’s unlikely to have neurons double counted between problems because 1) there were at least 15 days between the start of each new problem, and 2) the electrodes were advanced to obtain different populations of neurons after each problem.

Top of page 7 should read “while decoding at positions that were discriminable”. Also in that paragraph, “representing hidden task variables that in the first few days of training” has an extra “that”.

These typos have been corrected. Thank you!

Referee #3 (Remarks to the Author):

Summary

The authors present results from a study in which rats learned a complex task that required responding or withholding a response based on a sequence of presented odors. Rats learned four sequences; in two, decisions could be made purely based on the currently presented order, while in the other two sequences decisions had to take into account previous items in the sequence. After rats were well-trained on this task, they were presented with a set of 5 new problems with different odors while unit activity was recorded from neurons in lateral orbitofrontal cortex (OFC). As presumably different units were recorded on different days, the authors aligned activity for different problems by reducing the dimensionality of the population data and then projecting the data into a manifold that maximized correlation between two projected data sets. This procedure performed both for pairs of problems, to align data collected for a given animal, and for groups of rats. Within this aligned data, the authors found that during learning the population data in OFC came to distinguish between individual odors less. Instead, OFC developed a low-dimensional representation that distinguished between trials with different reward values. The dimensionality of the representations was greater for the more difficult sequences that required taking prior history into account. Trial type could be decoded across problems, suggesting that OFC formed similar representations of the task across problems. Similarly, trial type could be decoded across different rats after alignment, suggesting that different animals have similar representations in OFC.

Evaluation

There is substantial interest in neural coding of “schema” representations. This manuscript applies recently developed methods in a novel way to provide insights into how generalized representations are learned and represented in the brain. The study therefore addresses a question of broad interest. However, there are some limitations in the present manuscript that should be addressed. One central issue is that it is not precisely clear what information is contained within the OFC representations quantified in this manuscript. The manuscript would be improved if the authors could provide a more precise definition of schema and how the task structure is represented within the OFC schema. To what extent we should be surprised to find that cross-problem decoding works in this case, given the existence of task-sensitive activity in OFC for each individual problem?

Thank you for reading and commenting on our manuscript. We appreciate the importance of addressing shortcomings identified above. Below we will address them more specifically where they are raised under the Major Issues, however here we did want to briefly comment on the questions posed just above…..

Our working definition of a schema is that it consists of information about a problem or situation that is independent of the specific features of a given exemplar and thus can be used to facilitate learning/performance on other exemplars. An example would be the general layout of a city. Streets often form some sort of a grid pattern, tall buildings are at the center, residential areas farther out, and there often are ring roads going around the outside. We start to learn these general principles when we map our first city. This knowledge is not specific to that city however, and it forms a template that should make it easier for us to learn the specific layouts of new cities. Actually that last bit is not strictly true – it should help when cities obey these rules, but it could also hinder learning about a city that violates them…. but this latter interesting possibility is not one we explore here.

With regard to whether cross-problem/subject decoding is surprising; we believe that it is, in the sense that it did not have to be so. The neural networks might have developed a unique representation of each problem without extracting any generalizable structure, simply by pairing the specific odors (or odor sequences) with reward; just as knowing how to find an intersection in Baltimore does not help me find one in NYC, knowing that I respond after one odor but not after another in one problem does not help me know what to do when given new odors in another problem. Notably this possibility existed even if the animal were forming schemas, since generalized knowledge might not have been encoded in OFC. Or we might have uncovered intermediate results; for instance the only generalizable information could have been at P4 and P5 where odor identity is formally insufficient to predict reward. Another possible outcome was that the generalizable structure might have differed between subjects, so that cross-subject decoding would have been degraded. Or we might have found only very weak cross-problem/subject decoding.

Instead we found robust generalization of several different dimensions of information across both problems and subjects. We have tried to clarify our definition of schema and also add other information related to the above in the revision.

As no specific models of how the task schema might be represented are tested, it is hard to know how inevitable it is that different problems are represented similarly. For example, does OFC simply represent the reward on each individual trial, or does it also represent position and more abstract distinctions between diverging sequences? Is cross-problem similarity equally present for different aspects of the task (e.g., position, reward, sequence)? The manuscript focuses mainly on the reward dimension of the task structure. Without greater information about how OFC representations other task dimensions, it is not clear if OFC is truly representing a "schema" comprising multiple task dimensions.

As noted above, we now provide more concrete evidence regarding the information defining the schema in the form of the correlations between different task features and the CCs extracted by our analyses (**Figs. 3b-d** and **4b-d**; **Extended Data Figs. 9, 10, 12** and **13**). These new analyses show that the reward value of the trial is an important part of the schema code, but that aspects of the sequences are also prominently represented in other CCs. These new analyses show that the schema is multidimensional and provides labels for some of those dimensions. In addition, both sequence and position information that defines a shared task structure can be extracted through cross-problem and cross-subject decoding (**Extended Data Figs. 11** and **14**). We hope these new analyses address concerns that OFC just represents reward or that the information in the schema is not clear.

Major issues:

1. While evidence is presented that OFC represents a task schema that is common to different problems, it is less clear what information is represented in the schema that might drive performance. The results mainly focus on the distinction between rewarded and non-rewarded trials. What about other task-relevant information? For example, is the sequence (S1a, S1b, S2a, S2b) decodable? If so, at which positions? Does decoding generalize across positions? Does OFC represent the hidden structure in this task, or could it adequately be explained as representing only the reward at individual trials? There also appears to be position coding to some extent; characterization of this coding would help clarify not only the content of the schema but the role of OFC in representing different types of task content. Prior work has examined similar questions, but they are also relevant here. Particularly relevant is whether the different generalizable aspects of the task (i.e., features other than odor) are all transferred between problems, or whether only some aspects are represented in a generalizable way.

We agree that the information contained in the schema is of great interest, and we appreciate the reviewer prompting us to look more closely at this question. By definition, the cross-problem and cross-subject decoding is identifying information extracted and generalized across problems, so to address this question, we simply examined the relationship between the dimensions of the neural space containing the schema – as embodied in the CCs in our analysis – and a number of task features related to value and also to the sequences used, similar to the suggestions made above (**Figs. 3b-d** and **4b-d**; **Extended Data Figs. 9, 10, 12** and **13**). This approach revealed strong correlations between each of the first three components and task features; the first was best correlated with value, however the other two related most closely to aspects of the sequence itself. These new results, which essentially put a name to the information represented in the schema, are now presented in the main text. Additionally, we present a similar analysis for the other CCs in the supplemental figs (**Extended Data Figs 9, 10, 12,** and **13**). Lastly, we also analyzed our ability to decode sequences (**Extended Data Figs. 11a-b** and **14a-b**) and positions (**Extended Data Figs. 11c-f** and **14c-f**) across problems and subjects, as suggested explicitly above. We hope these analyses help address this important question.

2. The correlation between the neural measures and poke latency on each problem is unconvincing. There are only five observations and the variability between problems is almost entirely captured by two groups. The data would be better summarized as follows: compression, evolution, and behavior all increased after problem 1. Moreover, comparing the changes by trial within each problem would be more useful for testing for a relationship between brain and behavior.

We greatly appreciate this comment and suggestion, which we have followed in the revision by removing the correlations and instead focusing our analyses on the comparisons between Problem #1 and other problems. We hope this approach is more convincing.

3. The general approach used for the individual analysis techniques is generally well explained, but the details used for specific results is sometimes unclear. For example, I can guess at what the "pattern evolution" matrix represents (presumably a correlation between trial type dissimilarity matrices on each day and the day 15 matrix), but it is unclear from the text. In a few places, I was unclear what "data iterations" referred to.

We apologize for the lack of detail, and in the revision have tried to make the specific details of the data analysis and displays clearer. For instance, pattern evolution is now “schema evolution” and we refer to it where it is first derived in **Figs. 3g** and **4g**, before we use it in **Fig. 5d**. Hopefully that and the other changes will make things more understandable.

By “data iteration”, we meant “repeated analysis with a different pseudo-ensemble”. We ran most neural analyses 500 times. For example, the cross-problem decoding was repeated 500 times to obtain an averaged confusion matrix. For each time of repeated analysis, we generated a different pseudo-ensemble through subsampling and shuffling the trial order within each trial type (also to remove the temporal correlation between neurons within the same trial type). We deleted “iteration” and edited the text to better describe how we did the analyses. We have tried to clarify that in the methods, simply replacing this term with the term “repeat” or repeated.

Finally we have also now added tables in the Extended Data that give more specific details about each analyses, data used, and statistical results.

4. How did the manifold alignment between problem pairs allow for generalization to different problem pairs? If I understand the analysis, the manifold alignment seems very dependent on the ordering of the canonical correlations. The first three components seem well matched, which makes sense as the same dimensions may explain more variance for all problems, but how many other components were included in the classification analysis? How correlated were each of the component pairs? Analysis of the main components in terms of their sensitivity to task features and their correlation across problem pairs may help to address point 1.

Thank you for the excellent suggestion. In response, we examined the relationship between the CCs identified in the analysis and a number of task features, including those identified here and earlier. This revealed strong relationships between the initial CCs and some specific task features, which developed across training from Day 1 to Day 15 in both the cross-problem and cross-subject decoding analyses. These relationships are shown in **Figs. 3b-d**, **4b-d**, and **Extended Data Figs. 9, 10, 12** and **13**.

To directly answer the other questions here, we included all 60 CCs in the classifier even though many of them were not well matched between the training and test sets. To see the correlation between component pairs from the training and test sets, we plotted more CCs in the **Extended Data Figs. 9** and **12**. Whether these not-so-well aligned CCs were included or not had little impact on the decoding accuracy in these analyses. In considering this question, we think it is important to emphasize that our approach kept the training and test sets separate, meaning we didn’t use the correlation between CCs from the training and test sets to determine the number of CCs used in the classifier.

Other issues:

Figure 1j: Markers should be shown for all lines or none of them, and markers should be smaller if included.

We apologize for the confusion. The markers indicate statistical significance of the correlation. We have now added this detail to the figure caption.

Extended Data Figure 1: Reaction time should be split into correct and incorrect trials for the reward trials, or just correct trials included if there are not sufficient incorrect trials to estimate reliable statistics. For the non-reward trials, only incorrect trials should be shown. Currently, statistics for the non-reward trials are a mix of correct (2 s) trials and incorrect trials (varying time), which makes this figure hard to distinguish from the accuracy data shown in Figure 1.

We agree it is important to distinguish reaction time between the correct and incorrect trials, so we appreciate the reviewer's comment on this. The new analyses have been added to the **Extended Data Fig. 2** and **3**.

Why were the transition probabilities not equalized when transitioning between sequences? What determined those probabilities?

The rats could only be reliably expected to do only so many trials in a session. We did not want to compare rats doing different sequences, so we designed as random and counterbalanced a sequence as possible within that constraint. Once we mandated equal numbers of the 4 sequences and that the sequences not alternate, cluster, or clump together within the session, the transitions could not be fully counterbalanced. In the end, the ratios ended up a bit better than 60:40, so they are close to even but not perfect.

How did performance vary during days 15-23? Was the best day substantially better than the worst day?

All the rats were recorded for at least 15 days but not all rats were recorded for 23 days. To ensure roughly equal number of neurons on each day and stable performance on the final day, the data on Day 15 was taken from the day when the rats showed the best performance between days 15-23. Generally this day differed from the others only trivially and was never substantially better or worse than the overall trend, since performance was very stable once rats learned each problem. To illustrate this and to address any similar concerns readers may have as to the influence of this choice on our analysis, we now plot each rats' behavioral performance from Day 1 to Day 23 in the **Extended Data Fig. 1**.

The use of cross-ISI to select independent component runs should be explained further.

Thank you for the suggestion; we have added further explanation of the cross-ISI in the Methods.

Page 15: More detail should be given on the "template matching algorithm."

Thank you for the suggestion; we have added more information about the "template matching algorithm" in the Methods.

Page 15: Why were the electrodes advanced between problems to obtain different units?

Normally in our experiments we move electrode arrays daily and repeat recording on new problems, with acquisition of a new problem typically taking a single day. This approach is similar to what is conventionally done in unit recording experiments, where "fresh" neurons are isolated each day. This experiment was unique in that each problem took several weeks to learn. With many problems and rats involved, we anticipated that it would take 3-6 months to complete recording from each subject. This is too long to move the electrode arrays daily in this manner, but we worried that leaving them in one place for such a long time might result in unanticipated neural changes. So we adopted the compromise of moving them between problems. We felt that this was the best solution because it allowed us to maintain some connection to our prior data acquisition procedures. Importantly, moving between and not within the problems should bias against the observed results, since recording entirely different neurons on different problems should make cross-problem decoding less (and certainly not more) effective.

Page 17: What type of multidimensional scaling was used?

We used classical multidimensional scaling, also known as principal coordinates analysis, via the MATLAB function *cmdscale; w*e have added this information to the Methods.

I recommend representing the problems with colors that are distinguishable by colorblind individuals.

We apologize for not doing this and greatly appreciate the suggestion. In response, we have used Coblis, and online color blindness simulator, to adjust all of the color schemes to make all the figures friendly to colorblind individuals (e.g. protanopia, deuteranopia, tritanopia). We believe our new figures should be more easily distinguishable. In addition, we also added alternative features and labeling to improve readability. If what we have done is not sufficient, please let us know.

Reviewer Reports on the First Revision:

Referees' comments:

Referee #1 (Remarks to the Author):

The authors have satisfactorily addressed my prior concerns.

Referee #2 (Remarks to the Author):

Zhou et al. used an odor sequence task and numerous computational analyses on tetrode recordings in orbitofrontal cortex to study "schema formation" in rats. In their revision, Zhou et al. added the requested statistical tests, greatly improved the readability of the figures, and overall satisfactorily addressed my concerns. I am generally enthusiastic about this manuscript, I have a minor concern about the novelty of the findings. On lines 221-222 the authors describe their results as unique in finding neural generalization and faster behavioral learning. While this manuscript has many unique aspects, this particular finding is not completely novel. McKenzie et al., Neuron 2014 show this in their Fig 5.

See below for some additional minor concerns.

I would like the authors to discuss how behavioral output can influence evolution of OFC activity patterns (as opposed to vice versa). Stringer et al., 2019, Science show that brain regions reflect a combination of sensory input, internal states, and ongoing behavior. How might the present results be partially explained by an increase in behavioral stereotypy? An in-depth analysis is not necessary; a few sentences in the Discussion would suffice.

Fig 1h-j belong in Fig 5. There is hardly any discussion of this result or its implications until Fig 5. This will also help place into context the phrase "discounted future rewards", which is insufficiently explained in the text.

In the legend of Fig 2b, please add a + or – to indicate valence for the trial types like so: (Blue: S2a4+ and S2b5+, Red: S2a5- and S2b4-).

The legend of Fig 2d mentions error bars but they are not visible in the figure.

Please dedicate a short description of the task features described by the first 3 CCs. "Current Value" is intuitive, but "Odor-Unique Position" and "Alternating Position" are not. Perhaps a good alternative name for the former would be "Overlapping Odors" instead. In any case, it is not discussed in the text why "Alternating Position" would be an important feature to encode.

What is the rationale for specifically using the first 60 CCs (as opposed to any other number) in the Fig 3 and 4 analyses?

What do the authors mean by "meaning of underlying changes" on line 166?

"Externally available information" should just be "odor identity" on line 170 unless the authors have a specific reason for using that phrase. If so, that phrase is too vague.

"Hidden sequence information" should just be "previously presented odors" on line 171 for the same reason as above.

The phrases "external information" versus "hidden task variables" on lines 172-173 are also vague.

The text on lines 164-165 inadequately describe the analyses done in Fig 3g. "Decoding pattern" is so unclear that it's meaningless. Also it is not clear how the line plot on the right of Fig 3g was calculated. There should also be a statistical test to show that the values are increasing rather than simply showing that they are all above chance.

What is the rationale for Extended Fig 11e and 14e?

On lines 195-197 the clause starting with "an effect quantified by…" is way too long and too poorly worded to be meaningful. Please rephrase in simpler terms.

On line 198 "one" should be "on".

I am still unclear on what the "schema evolution" analysis is in Fig 5d. What is each data point? Please be explicit about what is being correlated.

In the paragraph beginning line 224 – "map" implies spatial. I recommend the use of the word "model" instead.

In Extended Data Fig 3 "negative trial type" should be "non-rewarded" to be consistent with previous terminology, unless "negative trial type" means something else.

In Extended Data Fig 5 "neural component" should be "canonical component" to prevent confusion.

Referee #3 (Remarks to the Author):

The authors were very responsive. In particular, the composition of schema representations in OFC have been clarified in the revised manuscript, which has improved the manuscript substantially. There are some remaining issues with the interpretation of links between neural representations of behavior and the interpretation of plots split by training and test data that should be addressed. I detail these concerns below.

1. Page 8: "This analysis showed that the pattern appeared more slowly on the first odor problem than one subsequent problems (Fig. 5b, Extended Data Fig. 15). Such accelerated learning is evidence of the operation of a schema. If the schema is manifest in the neural activity in OFC, then these behavioral changes should be mirrored in the development of the neural changes described above. Consistent with this idea, we found that both dimensionality reduction (Fig. 5c) and schema evolution (Fig. 5d) showed a similar acceleration between the first and subsequent problems."

Changes in behavior are shown within Day 1, while the neural measures are shown across days. The timescales are very different; thus, the link between these measures isn't evidence that behavioral changes are "mirrored" in the neural data. Presumably, if the correlation coefficients shown in Figure 5b were plotted over days instead of over trials, the curve would be mostly flat. There is evidence of a similar temporal scaling of learning in the S2a5- trials, which continue to improve over days and seem to accelerate somewhat after problem 1. However, curiously, the S2b4- trials seem to be learned more gradually after problem 1. Overall, there is evidence that learning accelerates for later problems, and that neural measures of dimensionality and schema evolution also accelerate for later problems. However, the relationship between the neural and behavioral measures is unclear. There are various reasons the two measures may not be perfectly linked, so this issue is not a major criticism. It does, however, need to be acknowledged and discussed in the conclusions.

2. Page 21: How were pairs of problems (for cross-problem classification) or pairs of rats (for cross-rat classification) selected? Were they randomly selected for each repeat? For the cross-problem classification, was selection randomly done separately for each rat (e.g., training might be problems 1 and 2 for one rat, and problems 1 and 5 for another rat), or were the same training and testing problems used across rats?

For the plots split by training and test sets in Figure 3, is the data averaged across rats? If so, that averaging will make the split between training and test for the cross-problem classification and cross-rat classification look similar almost as a matter of course. Also, is there any real difference between the training and testing data, given that they are (presumably) averaged over many repeats of the analysis? They would seem to be just different random samples of the same data. The finding that these random samples are more consistent on Day 15 than on Day 1 is still meaningful. However, that consistency might be partially driven by cross-subject consistency (as shown in Figure 4) rather than within-subject consistency. There may be a more direct way of measuring within-subject consistency, by calculating some measure of within-subject variance. Currently, if I understand correctly how they were calculated, the plots separating training and testing in Figures 3 and 4 may be largely redundant. One way to avoid some of these issues would be to collapse the training and testing sets in the Figure 3 plots; showing separate, random samples of the same data, averaged over many repeats, is not particularly meaningful. Then Figure 4 would just report the classification and schema evolution results. Another possibility would be to show an example of a single training and testing split within one rat in Figure 3, and one split across rats in Figure 4, though that may be noisy and not particularly representative. A final possibility would be to show some measure of mean variance between the training and testing sets, rather than plotting averages separately for the training and testing sets. The generalization between training and testing is still valid within the plots showing classification accuracy, as that is calculated within each repeat, but elsewhere it seems to be more misleading.

Minor issues:

Page 20: "we 1) randomly selected 140 neurons from each odor set" - Presumably "odor set" here refers to problem. If so, change to "problem" to make terminology more consistent.

Figure 5: "Trial type" isn't very descriptive and may easily be confused with the experimental design of the trials, which is also referred to as "trial type" (e.g., Figure 2a). I'd suggest instead using something like "discounted future reward" when referring to the conditions examined for the nose poke latencies analyses, so that trial type only ever refers to the experimental design.

Author Rebuttals to First Revision:

Note: Author rebuttals in blue

Thank you for the additional input. We have revised the manuscript to address each point. Our responses are detailed below point-by-point and major changes are highlighted in blue in the revised manuscript.

Referee #1 (Remarks to the Author):

The authors have satisfactorily addressed my prior concerns.

Referee #2 (Remarks to the Author):

Zhou et al. used an odor sequence task and numerous computational analyses on tetrode recordings in orbitofrontal cortex to study “schema formation” in rats. In their revision, Zhou et al. added the

requested statistical tests, greatly improved the readability of the figures, and overall satisfactorily addressed my concerns. I am generally enthusiastic about this manuscript; I have a minor concern about the novelty of the findings. On lines 221-222 the authors describe their results as unique in finding neural generalization and faster behavioral learning. While this manuscript has many unique aspects, this particular finding is not completely novel. McKenzie et al., Neuron 2014 show this in their Fig 5.

We apologize for this omission, and we now cite the paper and the implications of the results in Fig 5 in the second paragraph of the discussion.

See below for some additional minor concerns.

I would like the authors to discuss how behavioral output can influence evolution of OFC activity patterns (as opposed to vice versa). Stringer et al., 2019, Science show that brain regions reflect a combination of sensory input, internal states, and ongoing behavior. How might the present results be partially explained by an increase in behavioral stereotypy? An in-depth analysis is not necessary; a few sentences in the Discussion would suffice.

We appreciate this and have added a brief statement and the requested citation to the end of the second paragraph of the discussion making this point.

Fig 1h-j belong in Fig 5. There is hardly any discussion of this result or its implications until Fig 5. This will also help place into context the phrase “discounted future rewards”, which is insufficiently explained in the text.

We appreciate this suggestion however we would prefer to leave these panels in Fig 1, since they describe an important aspect of the behavior. Specifically this analysis shows that the rats attend to the sequence on positions other than P4 and P5, which is important for the bulk of the paper and not just Fig 5. The analysis in Fig 1h-j are also slightly different than those in Fig 5, since they collapse across problems. For these reasons we have left them in Fig 1, and instead changed our main text to make it more clear why these panels are relevant there. Of course if the reviewer feels they must be moved, we will be happy to make the change.

In the legend of Fig 2b, please add a + or – to indicate valence for the trial types like so: (Blue: S2a4+ and S2b5+, Red: S2a5- and S2b4-).

Done. Thank you!

The legend of Fig 2d mentions error bars but they are not visible in the figure.

The error bars were plotted in the figure; however they were so small as to be essentially invisible. To address this, we made the capsize (the length of caps at the end of error bars) bigger to aid the visualization of these error bars in Fig 2d.

Please dedicate a short description of the task features described by the first 3 CCs. “Current Value” is intuitive, but “Odor-Unique Position” and “Alternating Position” are not. Perhaps a good alternative name for the former would be “Overlapping Odors” instead. In any case, it is not discussed in the text why “Alternating Position” would be an important feature to encode.

We appreciate the suggestion to change these names to be more descriptive of these task features. In response, we have changed the names of the last two from “Odor-Unique Position” and “Alternating Position” to “Odor Overlap”, and “Positional Alternation”, respectively, and we have added text to more clearly describe them. “Positional Alternation” in particular is interesting since it resembles the periodic firing pattern exhibited by grid cell, thus its associated neural activity (CC #3) may be similarly relevant for providing basic task structural information that is invariant to the encoding of other two task features (“Current Value” and “Odor Overlap”).

What is the rationale for specifically using the first 60 CCs (as opposed to any other number) in the Fig 3 and 4 analyses?

There were 60 CCs in total for both training and test sets, and we used them all for the classification. For the training set in cross-problem classification, the CCA resulted in two paired matrices (480 trials x 30 CCs), corresponding to the neural data from two odor problems (or two rat groups in cross-subject classification). The two confusion matrices were concatenated into a single matrix (480 trials × 60 CCs), which was used as the training set. Using the same approach, we obtained the test set (480 trials × 60 CCs) from the other two problems. High correlations between CC pairs from the training and test sets are primarily restricted to the first a few CCs, however we chose to incorporate 60 CCs to provide a fuller representation of the data and to avoid selecting CCs based on their correlation between the training and test sets.

What do the authors mean by "meaning of underlying changes" on line 166?

This refers to the changes in the confusion matrix patterns, which reveal relationships between trial types by error decoding off the diagonal. We have changed the text to clarify this.

"Externally available information" should just be "odor identity" on line 170 unless the authors have a specific reason for using that phrase. If so, that phrase is too vague.

We have made the change suggested by the reviewer.

"Hidden sequence information" should just be "previously presented odors" on line 171 for the same reason as above.

We have made the change suggested by the reviewer.

The phrases "external information" versus "hidden task variables" on lines 172-173 are also vague.

We have deleted this sentence.

The text on lines 164-165 inadequately describe the analyses done in Fig 3g. "Decoding pattern" is so unclear that it's meaningless. Also it is not clear how the line plot on the right of Fig 3g was calculated. There should also be a statistical test to show that the values are increasing rather than simply showing that they are all above chance.

Decoding pattern refers to the pattern of decoding in the confusion matrices. This pattern reveals the detailed relationships between different trial types, much like a representational dissimilarity matrix (RDM). The correlation coefficient between two confusion matrices reveals the similarity between two corresponding "decoding patterns" or neural representations.

Based on this idea, we calculated the correlation coefficients between the two vectorized confusion matrixes obtained through cross-problem decoding on every pair of days (Day 1 – Day 15). The resultant 15x15 matrix was plotted as a heatmap on the left of Fig 3g. The line plot on the right of Fig 3g shows the last row in the heatmap with each data point indicating the correlation coefficient between the confusion matrix on each day (Day 1 – Day 15) and that on Day 15, which was taken as the final "template" pattern.

For clarity, we have now removed "decoding pattern" and simply refer to the pattern in the confusion matrices. We have also added text to describe what is correlated in Fig 3g more clearly. And we have added an arrow to show where the line plot on the right panel came from. Additionally, on the right panel of Figs 3g and 4g, we now use a z-test to test whether the correlation coefficient on day x (X-axis; Day 2 – Day 15) is significantly higher than that on a day that was y days lagged (Y-axis on the right; 1 – 14 days).

What is the rationale for Extended Fig 11e and 14e?

From cross-problem positional decoding analyses in Extended Fig 11e and 14e, we noticed that the improvement of cross-problem positional decoding mainly came from the first three positions. To confirm this impression, we repeated the cross-problem positional decoding but only with the first three positions. And the result agreed with our impression. As for why the positional decoding at the first three positions is exceptional, our speculation would be that at each one of the three positions,

the current and surrounding reward availabilities across sequences are similar, while at following positions (P4 and P5), the reward availabilities in S2b are not consistent with those in other sequences (S1a, S1b, S2a). We felt this additional analysis might be useful to remind readers about the differences between positions in the cross-problem positional decoding. We added this rationale to the figure legend.

On lines 195-197 the clause starting with “an effect quantified by…” is way too long and too poorly worded to be meaningful. Please rephrase in simpler terms.

We have rephrased the sentence.

On line 198 “one” should be “on”.

Corrected. Thank you!

I am still unclear on what the “schema evolution” analysis is in Fig 5d. What is each data point? Please be explicit about what is being correlated.

The schema evolution in Fig 5d used the same analysis as that in Fig 4g (and introduced in Fig 3g). The only difference is that in Fig 4g, the neural data from the 5 problems were combined, while in Fig 5d, the neural data from the 5 problems were analyzed separately.

In Fig 4, we performed cross-subject decoding. On each day, we obtained a confusion matrix, in which the trained classifier assigned each one of the 24 actual trial types with a predicted trial type label. The diagonal of the confusion matrix shows how well each trial type was successfully classified as itself, while the off-diagonal shows how each actual trial type was mis-classified as another trial type, which actually reflects relationships between trial types. Therefore the correlation between two confusion matrices tells the consistency or similarity between the underlying neural representations of all the 24 trial types.

Following this idea, we calculated the correlation coefficient between two confusion matrices obtained on different days to see how the confusion matrix across training days evolved to match the one on Day 15. The heatmap in Fig 4g shows the correlation coefficients of all day-pairs. When the confusion matrix (i.e., neural representation of 24 trial types) on Day 15 was treated as the final target of learning, we would get the line plots on the right of Fig 4g (i.e., the last row of the heatmap on the left).

We called the line plot in Fig 4g “schema evolution” because we extracted schema representation (as opposed to idiosyncratic representation) in the format of confusion matrix by cross-subject decoding; and the correlation between confusion matrix across days reflected an evolving process (i.e., becoming more and more similar) toward the target of learning.

We have changed the text to make this more explicit.

In the paragraph beginning line 224 – “map” implies spatial. I recommend the use of the word “model” instead.

We appreciate this concern. It is certainly true that the term cognitive map is often taken as synonymous with spatial map or “map of space in your head”. However we believe this is a distortion of the true meaning of the term. The original use of the term “cognitive map” by Tolman was to describe the causal relationships defining the animal’s environment, which is consistent with our usage of the term. Accordingly, although the term “cognitive map” has been used in the hippocampal field primarily to refer to spatial maps, it is widely though less famously used in other fields as an alternative term for the “cognitive model” in non-spatial domains. Nevertheless we are familiar with this confusion, so to try to address this we have modified this part of the discussion where we introduce, by saying “cognitive model” first, before we introduce the term map. We hope this changes is sufficient; since we connect our findings with an existing hypothesis (cognitive map hypothesis) in which OFC might be a part of the neural subtract important for cognitive mapping as defined by

Tolman and many in the field currently, we feel that moving fully to a new term would be even more confusing.

In Extended Data Fig 3 “negative trial type” should be “non-rewarded” to be consistent with previous terminology, unless “negative trial type” means something else.

We have made the change suggested by the reviewer.

In Extended Data Fig 5 “neural component” should be “canonical component” to prevent confusion.

We have made the change suggested by the reviewer.

Referee #3 (Remarks to the Author):
The authors were very responsive. In particular, the composition of schema representations in OFC have been clarified in the revised manuscript, which has improved the manuscript substantially. There are some remaining issues with the interpretation of links between neural representations of behavior and the interpretation of plots split by training and test data that should be addressed. I detail these concerns below.

We very much appreciated the reviewer’s additional in-depth comments, and we hope our responses will address these final concerns.

1. Page 8: "This analysis showed that the pattern appeared more slowly on the first odor problem than one subsequent problems (Fig. 5b, Extended Data Fig. 15). Such accelerated learning is evidence of the operation of a schema. If the schema is manifest in the neural activity in OFC, then these behavioral changes should be mirrored in the development of the neural changes described above. Consistent with this idea, we found that both dimensionality reduction (Fig. 5c) and schema evolution (Fig. 5d) showed a similar acceleration between the first and subsequent problems."

Changes in behavior are shown within Day 1, while the neural measures are shown across days. The timescales are very different; thus, the link between these measures isn't evidence that behavioral changes are "mirrored" in the neural data. Presumably, if the correlation coefficients shown in Figure 5b were plotted over days instead of over trials, the curve would be mostly flat. There is evidence of a similar temporal scaling of learning in the S2a5- trials, which continue to improve over days and seem to accelerate somewhat after problem 1. However, curiously, the S2b4- trials seem to be learned more gradually after problem 1. Overall, there is evidence that learning accelerates for later problems, and that neural measures of dimensionality and schema evolution also accelerate for later problems. However, the relationship between the neural and behavioral measures is unclear. There are various reasons the two measures may not be perfectly linked, so this issue is not a major criticism. It does, however, need to be acknowledged and discussed in the conclusions.

It is true that the behavioral changes did not perfectly mirror the neural data. Specifically the scales of the two measures differ. This reflects the resolution of the analyses; while we can analyze changes in behavior over just a few trials for each trial type within sessions, it is impossible to analyze changes in schema representation without a full session. So we cannot show neural effects at this time scale. However we of course think the two things are related, whether one causes the other or both are related to another process. In response we have added several sentences to make this mismatch clear to the reader and removed the term “mirrored” when discussing this relationship.

2. Page 21: How were pairs of problems (for cross-problem classification) or pairs of rats (for cross-rat classification) selected? Were they randomly selected for each repeat? For the cross-problem classification, was selection randomly done separately for each rat (e.g., training might be problems 1 and 2 for one rat, and problems 1 and 5 for another rat), or were the same training and testing problems used across rats?

For the plots split by training and test sets in Figure 3, is the data averaged across rats? If so, that averaging will make the split between training and test for the cross-problem classification and cross-

rat classification look similar almost as a matter of course. Also, is there any real difference between the training and testing data, given that they are (presumably) averaged over many repeats of the analysis? They would seem to be just different random samples of the same data. The finding that these random samples are more consistent on Day 15 than on Day 1 is still meaningful. However, that consistency might be partially driven by cross-subject consistency (as shown in Figure 4) rather than within-subject consistency. There may be a more direct way of measuring within-subject consistency, by calculating some measure of within-subject variance. Currently, if I understand correctly how they were calculated, the plots separating training and testing in Figures 3 and 4 may be largely redundant. One way to avoid some of these issues would be to collapse the training and testing sets in the Figure 3 plots; showing separate, random samples of the same data, averaged over many repeats, is not particularly meaningful. Then Figure 4 would just report the classification and schema evolution results. Another possibility would be to show an example of a single training and testing split within one rat in Figure 3, and one split across rats in Figure 4, though that may be noisy and not particularly representative. A final possibility would be to show some measure of mean variance between the training and testing sets, rather than plotting averages separately for the training and testing sets. The generalization between training and testing is still valid within the plots showing classification accuracy, as that is calculated within each repeat, but elsewhere it seems to be more misleading.

If we understand the reviewer correctly, this set of comments has two root concerns: (1) The first is that because both the training and test sets were randomly sampled from the same data, and the respective results were averaged across repeats, there should be no real difference between the two, thus it would be redundant to plot the analyses on training and test sets separately. (2) The second is that if training and test sets were not segregated by rat, then the apparent cross-problem classification might simply reflect cross-subject consistency.

Response to Concern (1):

The reviewer is concerned that separate analyses on the training and test sets would be redundant if they were averaged across repeats for the analyses in Figs 3 and 4. We appreciate this concern, however we think it does not affect the results in Figs 3/4b, which show only one example of a single split of training and testing, or in Figs 3/4 e-g, in which the classification accuracy was calculated separately on each repeat before averaging. We have clarified this in the legend. However, in Figs 3c-d, the results were indeed averaging across repeats for both training and test sets, and we agree that because of this, our strategy of separating training and test set data offers no real advantage. Thus, for these displays in the revision, we have followed the reviewer's suggestion and collapsed across the whole dataset, and relevant plots in Extended Data Figs 10 and 13 are also updated accordingly. Note this does not have any impact on the outcome of the analysis, since as the reviewer correctly points out, it is functionally the same as averaging across many splits.

Response to Concern (2):

The reviewer is concerned regarding how data were selected for repeats of the analysis in Fig 3 and that cross-problem classification is not within subject in this analysis. Here is how the data were randomly selected for each repeat. For each repeat of the cross-problem classification, the training and test sets came from different problems but always from the same rats (n = 9). For example, in one of the repeats, the training set might come from problems #1 and problem #2; and each one of the two problems combined neurons recorded from all the 9 rats. In other words, for the cross-problem classification, there was no chance that the training and test sets used neural data from different rats.

However it is also true that the training and test data in a given repeat included data from more than a single rat. This was necessary because individual subjects often did not provide enough neurons in individual sessions to permit a fully within-subject analysis across problems. To overcome this issue, we concatenated neurons recorded across all rats (n = 9) for each problem on each day. We used an assumption that neurons recorded from multiple rats can be seen as recorded from one single "virtual" rat. Based on this assumption, the training and test sets were always from the same "virtual" rat. We felt this was acceptable because it is a conventional approach and was the only way we could

analyze the data. Additionally it seems very unlikely that neural convergence would occur across-subjects and not within-subjects. Thus it seemed to us that the important question, once we demonstrated this convergence, was not whether some of it reflects within-subject consistency but rather whether any of it reflects across-subject consistency. This question led to our analysis presented in Fig 4, which isolates cross-subject classification. This analysis was not intended to be orthogonal to the analysis in Fig 3.

However we also think the two are not redundant, since in Fig 4, the neural data is combined across problems, so the classification might come from within-problem consistency. Figure 3 (cross-problem classification) rules this out. Or to put it more succinctly, Fig 3 shows across-problem consistency and Fig 4 shows across-subject consistency. We have now altered the section heading for these results and also edited the text to refer to across problem and across subject, removing specific references to within-subject. Additionally we have added several sentences to make clear that we did not isolate convergence within each subject although we presume that it occurs.

Minor issues:

Page 20: "we 1) randomly selected 140 neurons from each odor set" - Presumably "odor set" here refers to problem. If so, change to "problem" to make terminology more consistent.

We have made the change suggested by the reviewer.

Figure 5: "Trial type" isn't very descriptive and may easily be confused with the experimental design of the trials, which is also referred to as "trial type" (e.g., Figure 2a). I'd suggest instead using something like "discounted future reward" when referring to the conditions examined for the nose poke latencies analyses, so that trial type only ever refers to the experimental design.

We have made the change suggested by the reviewer.

Reviewer Reports on the Second Revision:

Referees' comments:

Referee #2 (Remarks to the Author):

The authors have satisfactorily addressed my concerns. Before recommending publication, I have just a few final points that the authors should address.

Because Nature has a wide readership, the authors should provide layman's descriptions of the following terms and what they represent:
- Manifold
- Manifold alignment
- Latent neural representation (particularly the "latent" part)

The authors' conjecture of the relationship between entorhinal grid cells and the positional alternation motif in CC #3 is unfounded. Unless the authors can produce both 1) a citation characterizing projections from entorhinal grid cells to OFC and 2) that grid cells can produce this periodic pattern between non-spatial trial types, that OFC might inherit, the authors should remove this statement. Instead, I think it is sufficient to speculate that this positional alternation pattern could arise simply from the trial structure. If the authors choose to accept this interpretation, they should additionally cite Bulkin et al. 2020, Hippocampus.

In the right-hand side panel of Fig. 3g and 4g, the Days Lagged axis is poorly described. Contrary to the legend, this data are not really "bars". Rather, they are filled squares if I am understanding it correctly. Additionally, the y ticks should have labels going from -1 to -14 days. I had missed this on the earlier version but the opposite y axis (Corr. Coeff.) should also have labels to indicate

the range of r values.

Referee #3 (Remarks to the Author):

The authors were very responsive to my most recent round of concerns, and the manuscript has been updated sufficiently to address them. I believe this manuscript would add substantially to the literature

Author Rebuttals to Second Revision:

Note: Author rebuttals in blue

Thank you for the additional input. We have revised the manuscript to address each point. Our responses are detailed below point-by-point.

Referee #2 (Remarks to the Author):

The authors have satisfactorily addressed my concerns. Before recommending publication, I have just a few final points that the authors should address.

Because Nature has a wide readership, the authors should provide layman's descriptions of the following terms and what they represent:
- Manifold
- Manifold alignment
- Latent neural representation (particularly the "latent" part)

We have modified the paper extensively to reduce it from 8.5 ~ 5 pages. As part of this, much of the introductory material where these terms were used was removed, and we now first introduce these concepts fully where we begin the analysis. At that point, we now explicitly define each of these terms in general or what we think are lay terms. We also cite the main study that (to the best of our knowledge) first applied this logic – and terms - to analyze information coding across sessions in which different neural subpopulations were likely sampled.

The authors' conjecture of the relationship between entorhinal grid cells and the positional alternation motif in CC #3 is unfounded. Unless the authors can produce both 1) a citation characterizing projections from entorhinal grid cells to OFC and 2) that grid cells can produce this periodic pattern between non-spatial trial types, that OFC might inherit, the authors should remove this statement. Instead, I think it is sufficient to speculate that this positional alternation pattern could arise simply from the trial structure. If the authors choose to accept this interpretation, they should additionally cite Bulkin et al. 2020, Hippocampus.

We have removed the speculation about grid cells. We did not add a citation, since we were already well above our limit of 30 citations, and we are not recording in hippocampus and the relationship is evident in the current data. We hope this is acceptable.

In the right-hand side panel of Fig. 3g and 4g, the Days Lagged axis is poorly described. Contrary to the legend, this data are not really "bars". Rather, they are filled squares if I am understanding it correctly. Additionally, the y ticks should have labels going from -1 to -14 days. I had missed this on the earlier version but the opposite y axis (Corr. Coeff.) should also have labels to indicate the range of r values.

Thanks for the suggestion; we have changed this so the description and labeling so the axes make more sense.