

Response to Review: Learning in visual regions as support for the bias in future value-driven choice

Sara Jahfari, Jan Theeuwes, Tomas Knapen

Dear Prof. Petersen,

Please find attached the revision of our manuscript entitled “Learning in visual regions as support for the bias in future value-driven choice” (CerCor-2019-00187) that my co-authors and I would like to resubmit for publication in Cerebral Cortex.

We thank both reviewers for their very nice words about the manuscript. Bellow you find a point-by-point reply to each of the reviewer’s comments. For clarity, the reviewer questions are written in *italic*.

We hope this revision adequately addresses the questions raised by both reviewers, and thank you for considering this MS for publication in Cerebral Cortex.

We look forward to your reply.

Sincerely,

Sara Jahfari (on behalf of both co-authors)

Response to Reviewer 1

In this paper, Jahfari et al present a reanalysis of BOLD data from a face variant of a well-established probabilistic reinforcement learning task in which participants are tested in extinction on the transfer of learned associations acquired during an initial learning stage. The focus of the paper is on the role visual areas (e.g. FFA and OC) in a visual corticostriatal loop (Seeger 2013) play in updating and deploying a representation of value during decision-making. This is an oft-disregarded topic in decision-making research, so this paper makes an interesting contribution to the field.

The paper was results rich and made use of an impressive array of methods, both on the behavioral model-fitting and fMRI side. I do however have some reservations regarding the interpretation of some key results.

1. The key analysis for the main argument of the paper is a decoding analysis which expands the set of classic value areas (e.g. striatum and vmPFC) and includes FFA and OC as ROIs. As I understand the RF approach, the classifier was trained on 2/3 of data in the transfer task to predict whether the participant would make the optimal choice or not. This classifier, is, in effect, learning to predict participant’s choices from brain data in the 9 regions of interest, and does so successfully on 69% of the trials in good learners.

+ I wasn’t sure, however, how much better one is doing by including the FFA. If FFA is critical for maintaining this value representation, then classification accuracy should be sensitive to the inclusion and exclusion of FFA.

The reviewer highlights an important point that we hope to clarify with this revision. The predictive accuracy of the RF classifier is not reduced with the removal of the FFA, or both FFA and OC. That is, we can just as accurately decode choice outcomes in the transfer phase by using only the single trial BOLD responses from striatal, prefrontal and motor regions (rFigure 1a). With the removal of just one or two regions (which notably can also be the putamen and caudate), the trial-by-trial data available to the algorithm from the remaining regions seems sufficient to reach the same predictive accuracy. At the same time, when evaluating

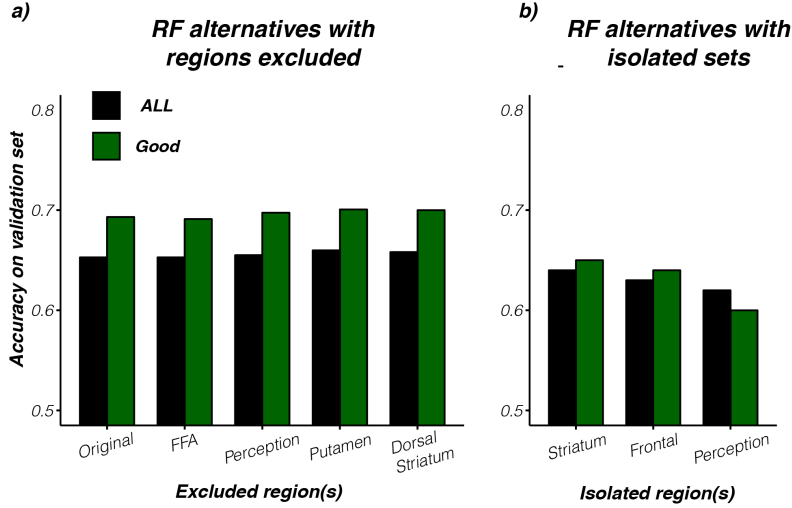


Figure 1: **Alternative evaluations with RF.** **a)** RF classification accuracy is not reduced with the removal of perceptual, or striatal regions when compared to the original model with 9 ROI's. Perception = exclusion of OC and FFA, Dorsal Striatum = exclusion of the Putamen and Caudate. **b)** However, accuracy is lower when striatal (Putamen, Caudate, and Accumbens), frontal (vmPFC, M1, DLPFC, and preSMA), or perceptual (FFA and OC) regions are evaluated in isolation. These alternative evaluations show that RF works best when trial-by-trial BOLD across multiple brain regions is combined.

data only from striatal, frontal, or perceptual regions we observe that the accuracy is reduced (rFigure 1b). This highlights that when the algorithm can combine information across regions predictions improve. These observations are now included in the results section and the supplementary Figure 5, with which we make explicit that the FFA is not crucial for the prediction of value-driven choice outcomes.

Our evaluation of both visual and traditional regions in predicting transfer phase choices was motivated by observations in the learning phase. There, we showed learning modulations of BOLD in both the FFA and OC by 1) $\Delta Value$ after the presentation of choice stimuli, and 2) RPE with feedback. In pursuit of understanding the purpose of these modulations for future decisions, we explored how these 'learning' visual regions would rank among more traditional regions - when the aim is to predict the outcome of a value-driven choice in the transfer phase. Our findings imply a supportive role for FFA in biasing the outcome of a value-driven choice because of 1) the observed BOLD modulations during learning, and 2) the high RF ranking of the FFA in transfer especially for good learners. In a task where participants pick the most valued face based on past learning and identification, our combined results imply a supportive role for the FFA in highlighting valuable face features.

+ Relatedly, Fig 6d says nothing about how much more important these regions are than regions which should not be expected to contribute to choice decoding at all (e.g. auditory cortex) — including some kind of negative control region would strengthen the interpretation of the results.

In response to the reviewers question, we now include evaluations with an additional random variable that was sampled from $\mathcal{N}(0, 1)$, and unrelated to the activity of any region or $\Delta Value$. Please notice that the random variable has negative importance in the ranking, meaning that removing it improves model performance. This additional analysis with the random control regions is now incorporated in the results section, Figure 6, and supplementary Figure 3.

2. The second comment concerns the neural interpretation of this decoding result. Given that the authors are analyzing data from a face task, this result seems consistent with the explanation that FFA is required to sustain attention to the task and disambiguate between perceptually similar faces. Would the authors expect

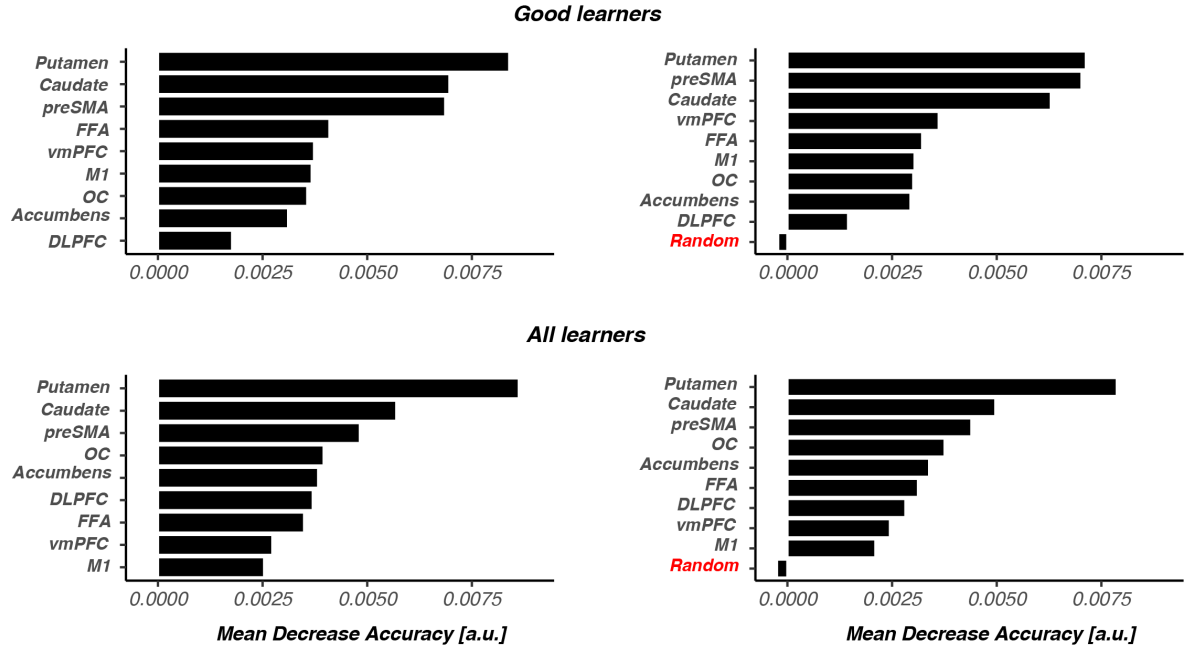


Figure 2: **Evaluation of RF ranking with a random control variable.** The left panel show the original ranking for the good learners (top) and all participants (bottom). On the right side, we evaluate ranking with all the original regions but now add a control region that was sampled from $\mathcal{N}(0, 1)$, and unrelated to the activity of any region or Δ Value. Notice that the random variable has negative importance in the ranking, meaning that removing it improves model performance with 0.5% (good learners) or 0.3% (all learners) points.

FFA involvement in a task with the same reward structure, but with different visual stimuli? (e.g. Frank et al 2004), or one in which selective attention to other visual inputs may be required (c.f. Leong et al 2017)?

We thank the reviewer for pointing us towards (Leong et al. 2017) paper. In response to the questions above we added the following section to the discussion.

“In the learning phase both OC and the FFA were modulated more by values of the (to be) chosen stimulus when belief representations were stable and distinct - i.e., we only observed differential Q -value modulations for the most reliable and easy to learn AB pair. This combined with the RPE modulations found in the same regions suggests an effect of value and learning on perceptual regions that is both specialized (FFA) and global (OC). Note however that this possibility must be studied further with designs that can zoom in on specificity with the separation of different perceptual dimensions (e.g., houses vs faces). Our transfer phase results imply a differential role for the specialized FFA, and the more low-level general OC, with the comparison of good vs all learners. Tasked with predicting the outcome of future value-driven choices RF rankings showed a specialized and prominent FFA role for good/efficient learners whereas OC was more important with the evaluation of all participants (where learning was less consistent or noisier across participants). Recent work on the interplay between learning and attention suggests a bi-directional relationship between learning and attention: we learn what to attend from feedback, and in turn, use selective attention to constrain learning towards relevant value dimensions (Rusch, Korn, and Gläscher 2017; Leong et al. 2017). In our study, better learning helps a more refined identification of rewarding features in a face, which we interpret as a narrower focus of selective attention in the FFA during learning (Niv et al. 2015). With past learning being more noisy, or less established, extraction of relevant features is less straightforward with attention being more spread to both specialized and global regions.”

3. Finally, were regions identified in the model-based analysis of the learning phase subsequently used as ROIs? Or were the same anatomical ROIs used for both the model-based and decoding analysis? It would be interesting, for instance, if the striatal areas identified as being modulated by value during the cue phase of learning (Fig. 4) are precisely the ones that, in concert with FFA, maintain a representation of value in the transfer phase. Or is the point simply to say that in value-based decisions, visual areas are modulated by the value of the chosen stimulus? (Fig. 4) If the latter, it might be helpful to compare with similar findings e.g. in Niv et al 2015.

We used the same anatomical ROIs for both the model-based deconvolution analysis (Figure 4&5) and the RF decoding analysis (Figure 2&6). In our approach, using the very same regions was essential to first explore *if* there is an interplay between learning and perception in a way that is comparable to traditional regions in the literature, to then proceed with the exploration of the *why*, or relevance, to future choices.

The methods section now better highlights that the same ROIs were used during training and transfer, and contains a better definition of how the anatomical masks were selected. We now also relate our findings to the Niv et al. 2015 study in the discussion (please also see our response to the previous question).

Minor comments:

Figures 1 and 6 should be combined somehow, as Fig 1. is a visual aid for understanding the analysis that pertains to results reported in 6.

Done. We now include graphics in figure 6 to illustrate how we relate the two phases of the learning task and the Q -learning analysis, to the RF analysis.

Why not do 3-fold cross validation? I.e. train on 2/3 of the data, test on 1/3, and iterate? That would be more convincing in light of potential overfitting issues.

In random forests, there is no need for cross-validation or a separate test set to get an unbiased estimate of the test set error. It is estimated internally with the use of the training set by using the out-of-bag

(OOB) samples. As with cross-validation, the classification error is computed using data that were not used for learning or tree construction (i.e., the OOB samples). In the MS, we now better explain this with the following in the method section:

“To achieve controlled variation, each decision tree is trained on a random subset of the variables (i.e. regions of interest chosen), and a bootstrapped sample of data points (i.e. trials or rows of the matrix in Figure 2c). In the construction of each tree about 1/3 of all trials is left out - termed as the “out-of-bag” sample – and later used to see how well each tree preforms on unseen data in the training set. Because in RF each tree is built from a different sample of the original data each observation is “out-of-bag” (OOB) for some of the trees. As such, each OOB sample is offered to all trees where the sample was not used for construction, and the average vote across those trees is taken as the classification outcome. The proportion of times that the classification outcome is not equal to the actual choice is averaged over all cases and represents the RF OOB error estimate. In other words, the generalized error for predictions is calculated by aggregating the prediction for every out-of-bag sample across all trees. In the results section, the OOB errors obtained from RF during training were well matched with the classification accuracy seen for the validation set given only the ‘good learners’ (OOB=30%, RF error validation set= 31%) or all participants (OOB= 33%, RF error validation set= 35%).”

The authors note that decoding only on data from “good” learners helps accuracy. This is not surprising. Participants who are more random, likely because they have not learned, would not be expected to maintain a value representation.

We agree with the reviewer. We reasoned that our RF predictions can result from alternative patterns in BOLD such as the build-up of a motor response, the ease of perceptual distinction between faces, or other to us unknown patterns. The evaluations comparing ‘good’ to ‘all’ learners, as well as, the relationship between $\Delta Value$ and RF uncertainty were control checks. Our aim was to ensure that our RF predictions are shaped by the consistency of past learning, and the representation of $\Delta Value$, prior to the evaluation of region importance. This is now better explained in the MS.

I would be curious to see a plot of the importance of FFA for decoding as a function of choicde difficulty, controlling for correctness — is it the case that for easy choices, participants have resolved the representation, and are thus engaging the FFA more?

We followed the reviewer’s suggestion and split the data into three separate sets based on ΔQ . This created an easy choice set - with $\Delta Value$ between the two choice options being large given beliefs that participants held at the end of learning - a medium choice set (not evaluated), and a very difficult choice set where $\Delta Value$ was very small (e.g., a choice between A:C, or B:D). RF was run separately for the easy and hard set with balanced numbers of correct and incorrect choices, where we again set aside 1/3 of each set for validation. Because the trial-by-trial BOLD data is cut in three this evaluation included all participants.

As plotted below, the pattern is opposite from what we would infer or was proposed by the reviewer. When $\Delta Value$ is large, RF predictions are best served by trial-by-trial BOLD from dorsal and ventral striatum, followed by vmPFC, the preSMA and M1. Here, regions involved with evidence accumulation (DLPFC), or perception (FFA and OC) rank last. Notably with easy choices the processing of BOLD from OC even has a negative effect on RF accuracy. Concurrently, with the evaluation of the most difficult choices - where participants decide between two very close in value positive (A or C) or negative (B or D) faces - we instead see processing of perceptual input to be far more important. When $\Delta Value$ is very small the caudate is followed by the FFA, OC, and preSMA in serving RF predictions. This suggests that only when value differentiation is most difficult past perceptual learning might serve the striatum by boosting the most rewarding face features.

We thank the reviewer for this important suggestion, which is now included in the supplement and incorporated in the MS results and discussion.

It’s unclear throughout what was predicted by the RF: after finally coming to the end I realized the classifier predicts whether the participant’s choice deviates from optimal or not, but the choice of words was somewhat confusing, and it seems simpler to just say, “we trained the classifier to predict the participant’s choice”

Done.

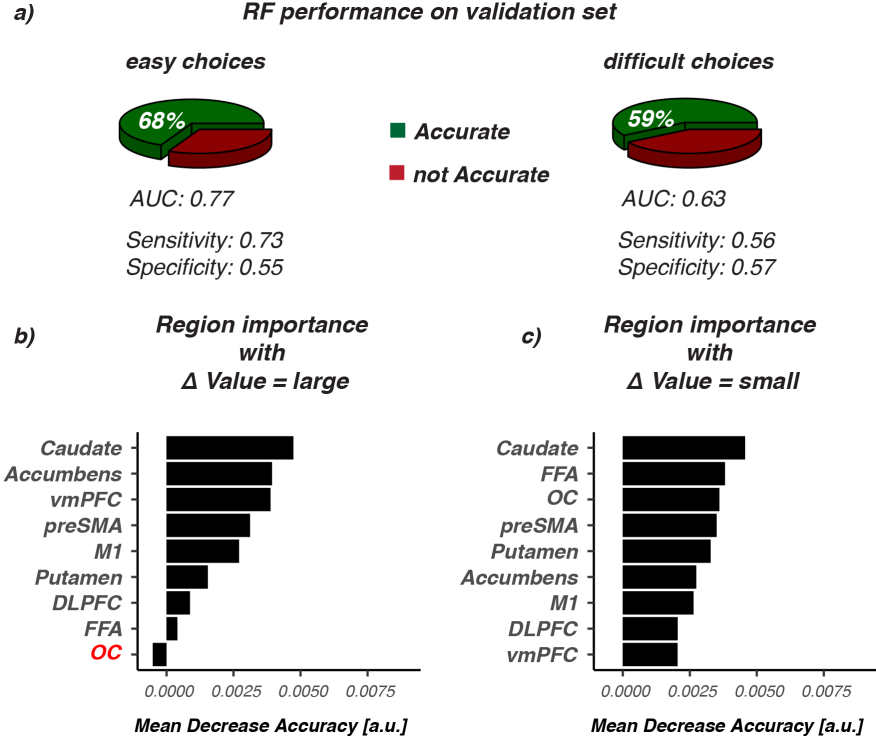


Figure 3: **Random Forest performance and ranking split for easy and difficult choice trials** . For this analysis we created an easy choice set - with Δ Value between the two choice options being large given participants beliefs at the end of learning - a medium choice set (not evaluated), and a very difficult choice set where Δ Value was very small. RF was run separately for the easy and hard set with balanced numbers of correct and incorrect choices, where we again set aside 1/3 of each set for validation. Because the trial-by-trial BOLD data is cut in three this evaluation included all participants (N=43). **a)** Shows the classification or decoding accuracy (green) given the separate unseen validation sets for ‘easy’ value-driven choices where Δ Value is large (e.g., a choice between A and D, or C and B), or ‘hard’ choices with small Δ Values (e.g., between A and C, or B and D). **b)** Plotted ranking of the ROIs from the RF model for easy, or hard choices **(c)**. Notice that with easy choices, OC has a negative ranking. Removing OC here improves RF decoding.

The last para on page 16 should be unpacked a bit for an audience naive to RF classification. What is chance in Fig. 6d?

This part of the results section is now better explained for naive readers, and extended with the additional control checks suggested above.

Fig 6 caption has a typo: “Each tree is build” should be “each tree is built”

Corrected.

Response to Reviewer 2

I'd like to complement the authors on a very well written paper and well executed work. Understanding how perceptual and value processing interact during value-driven choice is a very timely topic. I would like to see some additional data analysis using random forest procedure (RFP).

1. In RFP, correlated features will show in a tree similar and lowered importance, compared to what their importance would be if the tree was built without correlated counterparts. Prior literature reported significant functional connectivity between FFA and striatal regions. Thus, I would expect that predictors used in the analyses do significantly correlate. Was it so in your data?

Yes, this is an important point that we now better explain. Based on our earlier work with this data focusing on connectivity we were aware that the single trial BOLD across these regions was correlated. To make this explicit and to explain how we dealt with the problem of co-variation we now include:

1. The correlation matrix below with the single trial BOLD relationships across regions, or with Δ Value.
2. An elaboration on how RF deals with colinearity in predictions.

“With RF the problem of correlated features is minimized for predictions with variable selection - i.e., the random selection of a set of regions to use for each tree. With more variables selected, we get better splits in each tree but also highly correlated decisions trees across the forest, which diminishes the forest effect. To find the best balance, this study optimized the number of variables to select with a tuning function using the OOB error estimate.”

3. And an explanation of how we dealt with the problem of colinearity in the evaluation of ranking.

“While potential confounds of colinearity on the RF ranking cannot be excluded, we tried to minimize this with the use of permutation importance. Here, by using the OOB samples the importance of each variable (region) is computed as the difference between the models baseline accuracy and the drop in overall accuracy caused by permuting that column (region). While being more slow, permutation importance is described as more robust in comparison to the default (gini) importance computation where only the uncertainty of predictions is evaluated (with no checks on accuracy fluctuations after region permutation).”

Please note, that although we were aware of the correlations between our variables with the initial submission and therefore used permutation importance from the beginning, further research into RF ranking for this revision suggested that it is best not to scale the RF output (for example see <https://explained.ai/rf-importance/index.html>). As we were previously unaware of this detail, we ran all our initial RF analysis anew with scaling set to False for permutation importance ranking (this includes all additional revision analysis). The new ranking was identical evaluating all participants. With good learners, the ranking was identical in all regions but the accumbens and OC in the tail. That is, the top 5 regions were the same, but in the tail the accumbens switched places with OC. We further note, that these new evaluations were

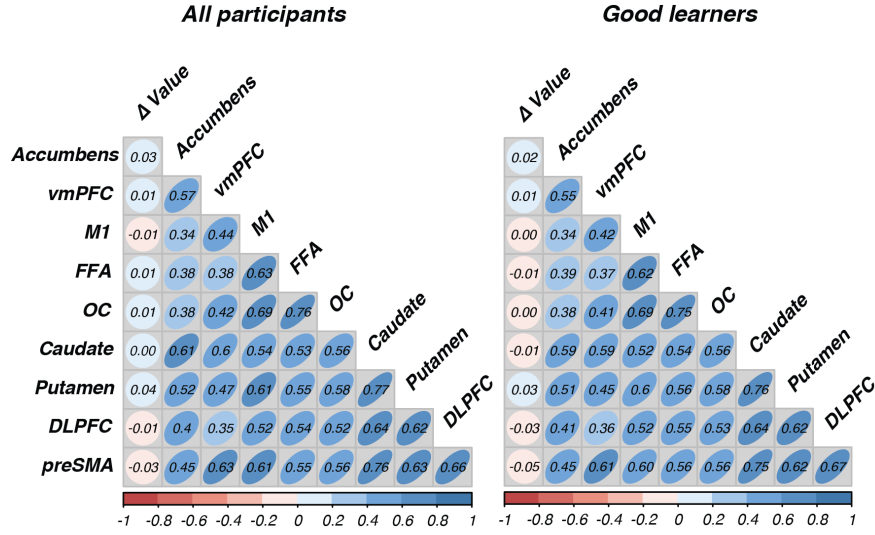


Figure 4: **Correlation Matrix for single trial BOLD across regions, or with Δ Value.** The left panel evaluates all participants (N=43), with the good learners (N=34) on the right. Centroids combine shape and strength of the relationship with the number on top of each box being the correlation value. It stands out that across all brain regions trial-by-trial BOLD is correlated, where we see the strongest relationships between regions of the dorsal striatum (i.e., Caudate and Putamen), visual cortex (i.e., OC and FFA), or the PFC and the dorsal striatum (i.e., Caudate and preSMA or DLPFC). These observations concur with our connectivity evaluations of this data showing effective connectivity inputs from the preSMA and DLPFC into the dorsal striatum, and earlier work reporting functional and effective connectivity between the FFA and striatum. Also note the lack of relationship between Δ Value single trial BOLD from any region. This contrasts the relationship between Δ Value and RF predictions, where the latter combines trial-by-trial information across regions to predict a choice.

performed with a newer version of the randomForest package in R (randomForest_4.6-14), which slightly improved our classification accuracy to 70% (was 69% in initial submission) for the good learners.

+ If it was, it might better to select features recursively, rather than all together. This will also allow contrasting prediction accuracy with and without visual regions. If you find that including visual regions significantly improves accuracy, it would strengthen your results (in my opinion).

Our purpose with RF, in the transfer phase, was to evaluate how the ‘learning’ perceptual regions would rank among more traditional regions in the prediction of choice. The permutation importance used evaluates changes in the full model baseline accuracy with the permutation of each predictive region. As such, ranking is based on the anew evaluation of accuracy with each region randomized. Because our top 3-ranking regions were correlated more than most other regions, recursive feature selection would unfortunately not clarify problems with colinearity or help the interpretation of ranking.

We hope that our response to reviewer 1 (first question) sufficiently explains how the predictive accuracy of RF does not decline with the removal of just one or two perceptual, or striatal regions. This, in our interpretation is related to the correlation pattern across the RF regions, potentially caused by general learning modulations in the learning phase. Consistently, when RF is given random input that is unrelated to learning, we observe that accuracy is decreased (please see Figure 6d and supplementary Figure 3).

Other suggestions are mostly cosmetic:

on p.8 some abbriviations show up for the first time without being defined: PPA, LOC.

Defined. Thank you for pointing us to this.

Please describe how you define masks for vmPFC, DLPFC, preSMA, and M1.

The following was added to the methods section of the MS.

“The DLPFC template was obtained from an earlier study, linking especially the posterior part to action execution (Cieslik et al. 2012). The preSMA, vmPFC, and M1 templates were created from cortical atlases available in FSL. All anatomical masks, and the localizer group-level FFA mask can be downloaded from github (see acknowledgements).”

Emergent literature on the role of visual attention in a value-based choice suggests that the relation is bidirectional. It might be relevant to include this point in the introduction/discussion

Following this suggestion from both reviewers this is now included in our discussion section. For the full text that was included we kindly point the reviewer to our response on point 2 from reviewer 1.

Greatly enjoyed reading about your work!

We thank the reviewer for the encouraging words and questions in the review.

References

- Cieslik, E C, K Zilles, S Caspers, C Roski, T S Kellermann, O Jakobs, R Langner, A R Laird, P T Fox, and S B Eickhoff. 2012. “Is There ‘One’ Dlpfc in Cognitive Action Control? Evidence for Heterogeneity from Co-Activation-Based Parcellation.” *Cerebral Cortex* 23: 2677–89.
- Leong, Yuan Chang, Angela Radulescu, Reka Daniel, Vivian DeWoskin, and Yael Niv. 2017. “Dynamic Interaction Between Reinforcement Learning and Attention in Multidimensional Environments.” *Neuron* 93: 451–63.

Niv, Y, R Daniel, A Geana, S J Gershman, Y C Leong, A Radulescu, and R C Wilson. 2015. “Reinforcement Learning in Multidimensional Environments Relies on Attention Mechanisms.” *Journal of Neuroscience* 35: 8145–57.

Rusch, Tessa, Christoph W Korn, and Jan Gläscher. 2017. “A Two-Way Street Between Attention and Learning.” *Neuron* 93: 256–58.