

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

Sara Jahfari, Jan Theeuwes, Tomas Knapen

5/20/2019

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 (Seger 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 (SFigure 1a). With the removal of just one or two regions (which

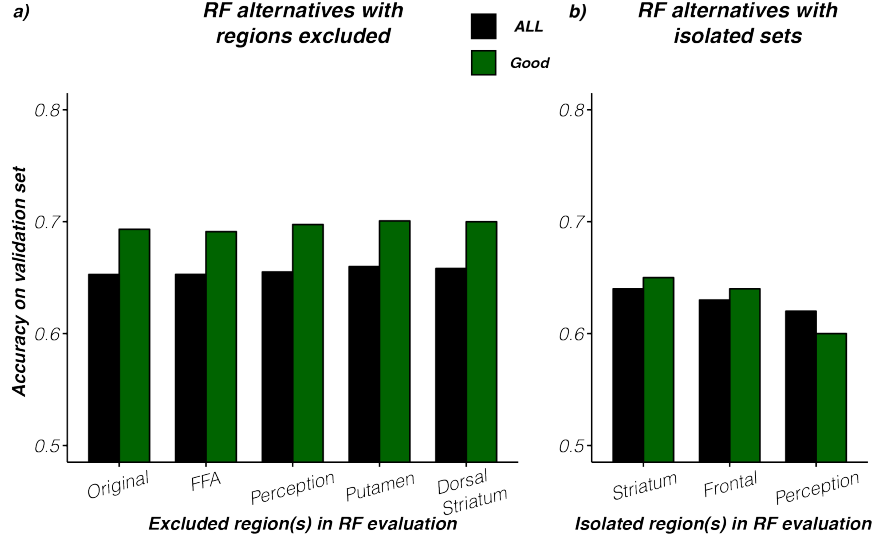


Figure 1: **Alternative evaluations with RF.** **a)** RF classification accuracy is not reduced with the removal of perceptual, or striatal regions when compared to the original model with 9 ROI's. Perception = exclusion of OC and FFA, Dorsal Striatum = exclusion of the Putamen and Caudate. **b)** Accuracy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

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. However, with the evaluation of data only from striatal, frontal, or perceptual regions we observe that the accuracy is reduced (SFigure 1b). This, highlights that when the algorithm can combine information across regions predictions improve. These observations are now included in the results section on page xx and the supplementary and should make it 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 in the future, 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. We suggest our findings to 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. Hence, in a task where participants pick the most valued face based on past learning, 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.

Unfortunately, we could not evaluate or add auditory cortex to our RF model as a control because of dropout artifacts (near the ears). We previously only interpreted the ranking of the highest regions, and perceptual regions. To be more complete, we now better integrate the learning phase results with the RF ranking and additionally focus on providing an interpretation for the lowest (DLPFC) ranking region. Additionally, we now make it explicit that our interpretation of ranking is limited by not having a control region that is untied to decision-making or learning processes.

The following was added to page xx, in the MS.

The DLPFC ranked lowest for good learners during transfer, where it during the learning phase was not reliably modulated by either differential value or RPEs. With the evaluation of all participants however this RF ranking was higher. BOLD responses in the DLPFC associate with evidence ambiguity, evidence accumulation, and, the integration of perceptual information into choice evidence [e.g., @Bachetal2009; @Kahntetal2011; @Piliastidesetal2011]. Additionally, a recent study, focusing on how dimensions are reduced for selective attention in learning, reported a fronto-parietal network including the DLPFC when participants have not yet learned the reward relevant features [@Nivetal2015]. With less consistent learning the DLPFC up-ranking could result from more ambiguity [@Bachetal2009], the upholding of multiple stimulus features with the employment of a broad fronto-parietal network [@Nivetal2015], or a bigger dorsal striatum reliance on the DLPFC for inputs in support of the perceptual choice [@GoldShadlen2007; @Heekerenetal2008]. For good learners, top-down DLPFC inputs can be bypassed with differential value computations and directed selective attention to high value face features by the FFA - which connects directly to the striatum [@Jahfarietal2015; @Kravitzetal2013]. Future studies need to confirm our RF ranking observations with the evaluation of a control region that is indifferent to decision-making or learning, or, by using combined decision-making and learning theories to separate value from evidence accumulation [@Pedersenetal2017; @Fontanesietal2019].

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 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. We have added the following section to the discussion in page xx in response to answer the reviewers questions above.

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 together 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). This possibility however should be studied further with designs that can zoom in on specificity with the separation of different perceptual dimensions (e.g., think of houses vs faces). While our design, with only faces, precludes any direct claim on specificity our analysis of the transfer phase does show a differential role for the specialized FFA and the more low-level general OC in the evaluation of good vs all learners. Tasked with predicting the outcome of future value-driven choices the RF rankings show a specialized and prominent FFA role for good/efficient learners whereas OC was more important with the evaluation of all participants (where learning was more consistent or noisy 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 [@Ruschetal2017; @Leongetal2017]. 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. When learning is more uncertain, the 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 (Figure 4&5) and decoding analysis (Figure 6). In our approach, the use of 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.

In isolation, no single node in the brain can achieve the goal of deliberately initiating a value-driven or voluntary choice. Indeed, we found RF predictions to decline when striatal, frontal, or visual regions are evaluated in isolation. The RF evaluations in the transfer phase rank the two dorsal striatum regions that were sensitive to both differential value up-to a choice (Figure 4) and RPEs (Figure 5) as most important for the prediction of future value-driven choices. The high ranking of the FFA with established learning implies a role for the FFA in supporting the striatum gate the most valued choice. This concurs with theories where learning narrows selective attention (Niv et al. 2015; Leong et al. 2017), our effective connectivity work showing the involvement of a direct-indirect fronto-basal ganglia network in value-driven choices (Jahfari et al. 2018), and direct effective connectivity inputs from the FFA into the dorsal striatum during perceptual decisions about a face (Jahfari et al. 2015).

Bellow, in response to reviewer 2 (question xx) we elaborate on how the anatomical regions were selected.

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.

over nadenken

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 highly 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 could have resulted from alternative patterns in BOLD such as the buildup of a motor response, the ease of perceptual distinction between faces,

or other to us unknown patterns. The additional 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 RF predictions are shaped by the consistency of past learning, and the representation of $\Delta Value$, before moving on 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 reviewers 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 bellow, the pattern is opposite from what we would infer or was proposed by the reviewer. When $\Delta Value$ is big, RF predictions are best served by trial-by-trial BOLD from both the ventral and dorsal striatum, followed by M1, and the vmPFC. Here, regions involved with the processing of choice difficulty or conflict (preSMA), perception (OC and FFA), and evidence accumulation (DLPFC) rank last. 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 and conflict to become 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 when value differentiation is most difficult the specificity of past perceptual learning might serve the striatum by boosting the most rewarding face features.

We thank the reviewer for suggesting this additional analysis (now included in the supplement), which helped us refine our interpretation in what role perceptual regions play in the RF predictions of future value-base decisions.

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.

The following add this to results section (287).

The RF classifier was trained to predict the participant’s choice, on each trial, given trial-by-trial BOLD estimates from striatal, prefrontal, and viusal regions.

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

Besides providing insights into how BOLD responses in the transfer-phase contribute to predict value-driven choice outcomes (i.e., whether participants would choose the option with the highest value given past learning) the RF algorithm additionally outputs a hierarchy, thereby ranking the contribution of each region in the achieved classification accuracy. Figure 6d shows the ranking of all ROIs for good learners where the model had the highest predictive accuracy. First, regions in the dorsal striatum were most important, which aligned well with both the literature and the BOLD modulations we found by $\Delta Value$ and RPE during the learning phase. These regions were next followed by the preSMA. Evaluation of this region during the learning phase showed no modulations by $\Delta Value$ or RPE on BOLD (supplementary Figure 1&2). Nevertheless, this region is typically associated with choice difficulty/conflict and might be essential in the resolve of a choice when value differences are small. Remarkably, the third region in this hierarchy was the FFA. In a task where participants pick the most valued face based on past learning, this ranking of the FFA just above the vmPFC and accumbens (ventral striatum) implies that the $\Delta Value$ and RPE modulations of BOLD observed during learning could function to strengthen the recognition of valuable features. Note, however, that with the evaluation of all participants – including some who were less good in learning – the ranking of both the FFA

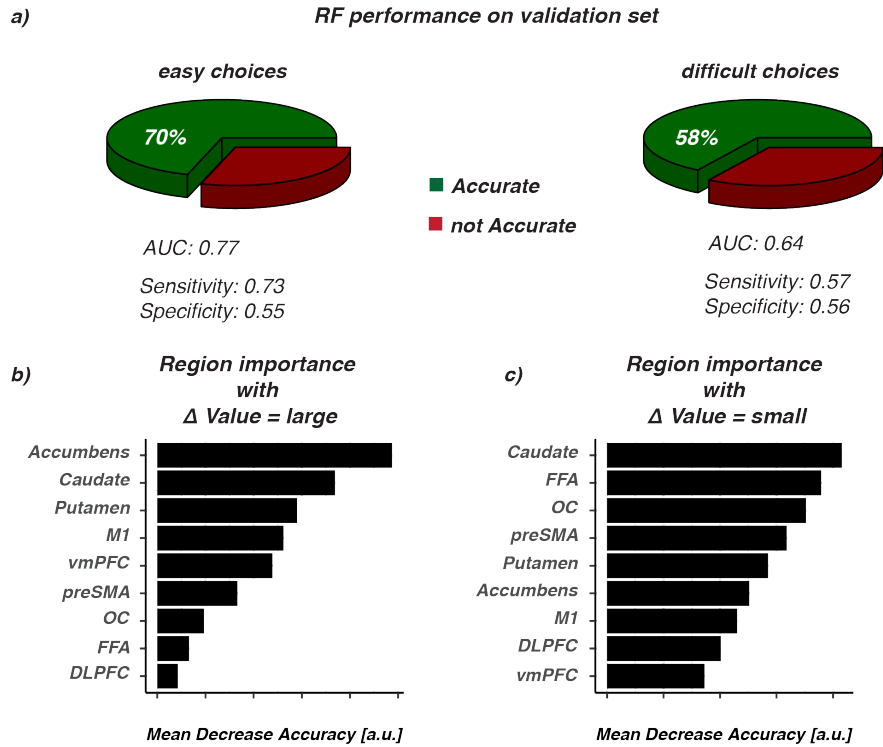


Figure 2: **Random Forest performance and ranking split for easy and difficult choice trials** . a) Shows the classification or decoding accuracy (green) given the separate unseen validation sets for ‘easy’ value-driven choices where Δ Value is big (e.g., a choice between A and D, or C and B), or ‘hard’ choices with small Δ Value (e.g., a choice between A and C, or B and D). b) Plotted ranking of the ROIs in their contribution to the predictive accuracy of the RF model for easy, or hard choices (c). Because of the large reduction in trial for this split, the RF evaluation includes all participants (N=43).

and vmPFC was much lower (please see supplementary Figure 3b), which might be caused by more noise across the group in learning. We will return to this point in the discussion.

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

Done.

Response 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?

[Analyzes] kijk naar de correlatie matrix single trial bold. [theorethisch, correlatie in RF en hoe het daarmee omgaat]

We hebben inderdaad zelf ook in eerder studies laten zien dat er een relatie bestaat tussen het BOLD response van de FFA met de striatum. Met name voor dorsaal, en veel lager voor ventraal.

If it was, it might better to select features recursively, rather than all together.

kijk hier even naar: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-8-25#Sec12>

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).

[Analyzes] kijk naar model met en zonder visuele gebieden. Volgens mij maakt dit voor accuracy weinig uit. Leg uit waarom je approach wel goed vindt. Refereer ook naar R1 met het toevoegen van niet belangrijke gebieden. Laat misschien ook zien dat voor geen gebied single-trial BOLD correleert met delta value. Dus wat de classifier oppakt van de gecombineerde data is gerelateerd aan delta value en keuze uitkomst.

Other suggestions are mostly cosmetic:

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

[praktisch] corrigeer, en bedank

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

[praktisch] voeg toe

The DLPFC template was obtained from an earlier study, linking especially the posterior part to action execution (Cieslik et al., 2013).

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

[theoretisch] zoek dit op en weef het in de introductie/discussie

Greatly enjoyed reading about your work! Thank you!

References

- Jahfari, S, K R Ridderinkhof, A G E Collins, T Knapen, L J Waldorp, and M J Frank. 2018. “Cross-Task Contributions of Frontobasal Ganglia Circuitry in Response Inhibition and Conflict-Induced Slowing.” *Cerebral Cortex*, bhy076.
- Jahfari, S, L Waldorp, K R Ridderinkhof, and H S Scholte. 2015. “Visual Information Shapes the Dynamics of Corticobasal Ganglia Pathways During Response Selection and Inhibition.” *Journal of Cognitive Neuroscience* 27: 1344–59.
- 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.