**Response to Reviewers**

Large-scale meta-analysis of human medial frontal cortex reveals tripartite functional organization

Alejandro de la Vega, Luke J. Chang, Marie T. Banich, Tor D. Wager & Tal Yarkoni

Dear Dr. Shaham,

Thank you for the opportunity to revise our manuscript by responding to the reviewers’ insightful comments. Below we outline the major changes to the manuscript to address the concerns by the reviewers, followed a point-by-point response to various specific concerns.

1. We have revised our introduction to more clearly motivate meta-analytic co-activation based clustering. We have emphasized that morphological and connectivity based parcellations infer functional divisions; whereas co-activation based clustering more directly identifies regions with similar co-activation across a wide variety of psychological manipulations.
2. We have more clearly distinguished in our results where our parcellation coincides with, or differs from, previous work, and have used the Harvard-Oxford probabilistic atlas to generate more accurate labels.
3. We have simplified the terminology, more liberally cited Figure 1, and revised the methods section to improve clarity.
4. We added post-hoc exploratory tests to determine if certain topics were more strongly associated with particular regions. These post-hoc tests allowed us to address Reviewer 1’s concerns about differences within the middle zone, and Reviewer 2’s question about function gradients.
5. We have updated Figure 2 to include results from 12 clusters and display the results from 9 clusters in more detail using coronal slices.
6. The discussion has been substantially revised. In particular, we have been more careful in describing the limitations of our work, and have expanded the discussion of potential future challenges.

We believe that the feedback provided by the reviewers has substantially improved the manuscript, and hope that the extensive changes have addressed all of the reviewers’ concerns.

Sincerely,

Alejandro de la Vega

All Reviewer comments are denoted in italics. Quotes from our revised manuscript are italicized and underlined.

**Reviewer 1**

Because of the length of this review (which we found very helpful!), we have not explicitly reproduced all of the reviewer’s comments. Instead, we have summarized our responses to the reviewer’s major suggestions. We have also included a point-by-point response to minor comments not encompassed by the major suggestions. We do not explicitly address the typos or grammatical changes suggested by the reviewer, but have followed the reviewer’s recommendations in virtually all cases.

**Major / General**

**Introduction**

*“The rationale for partitioning MFC on the basis of meta-analytic co-activation is inadequate, given that this is the central means of identifying the parcels for subsequent profiling*

We have addressed all of the reviewer’s specific comments through extensive revisions to the introduction. We now open the introduction more broadly, and establish the problem before jumping into the various fMRI associations with MFC. We also more carefully outline the limitations of morphological, cytoarchitechtonic and connectivity based parcellation methods, and highlight the advantages of a co-activation based clustering in relation to previous approaches. Major changes can be found on pp. 4-8; here we highlight a few key sections:

*“The medial frontal cortex (MFC) plays a key role in contemporary accounts of various psychological processes, including motor function, cognitive control, emotion, pain and social cognition. However, the precise correspondence of psychological states onto discrete medial frontal anatomy remains elusive. There have been several recent attempts to define distinct sub-regions of MFC by inferring functional differences on the basis of morphology (Vorobiev et al., 1998; Vogt, 2009; Palomero-Gallagher et al., 2013) and in-vivo structural (Beckmann et al., 2009; Sallet et al., 2013; Neubert et al., 2015) and functional connectivity (Andrews Hanna et al., 2010). Although these studies have greatly advanced our understanding of MFC’s organization, such methods cannot directly demonstrate the functional significance of the resulting regions as they do not measure their response to various psychological challenges.” – lines 48-57*

*“Because most researchers tend to be intimately familiar with a particular psychological domain, most meta-analyses are restricted to a relatively small subset of empirical findings. Even those meta-analyses that take a broader look at organization of the MFC typically only include a subset of cognitive states hypothesized to be important (e.g. negative affect, pain cognitive control; (Shackman et al., 2011) or restrict themselves to a small region of interest (e.g. (Palomero-Gallagher et al., 2015). This narrow scope necessarily limits the ability to address the specificity of activation of psychological states across MFC anatomy. That is, without considering a wide representative range of psychological states, it is difficult to determine whether particular kinds of tasks preferentially recruit the MFC.” – lines 71-79*

**Methods & Results**

*“a. Yes, but at times the complex methodology is challenging to follow. Specifically, the**terminology is unwieldy at points (voxels, features, parcels, zones, sub-regions, and ROI's); be consistent.”*

“*fig 2 - why is it zones on the left and sub-regions on the right?”*

We have revised the terminology used throughout the manuscript to ensure clarity and consistency. We now avoid the terms “features” and “parcels”, instead referring to “topics” and “clusters”, to be consistent. We also now explicitly indicate that we will refer to clusters in the 3-clusters solution as “zones” and those in the 9-cluster solution as “sub-regions”. Finally, we consistently use the term “psychological topics” instead of “concepts”, “functions” or other such terms.

*“We henceforth refer to the clusters from the 3-cluster solution as “zones” to differentiate them from clusters in the 9-cluster solution, which we refer to as “sub-regions”. – lines 262-264*

*b. Figure 1 is incredibly helpful, yet, is not cited in the text. It would be helpful to liberally cite each panel as you work your way thru the constituent methods*

We now refer to Figure 1 liberally throughout the Methods, making sure to reference every panel of Figure 1 as we proceed with our description.

*“c Not clear whether co-activation is w/in or b/w studies. If between, why?”*

We have expanded our description of the co-activation clustering methodology to indicate that we are calculating the correlation *across* studies between MFC voxels and whole-brain PCA components.

*“Next, we calculated to co-activation of each MFC voxel with the rest of the brain by correlating their activation pattern across studies with the rest of the brain. Activation for each voxel is represented as a binomial vector of length 9,721, with 1 indicating that that voxel was active in a particular study. As correlating the activation of every MFC voxel with every other voxel in the brain would result in a very large matrix (15,259 MFC voxels x 228,453 whole-brain voxels) that would be computationally costly to cluster, we reduced the dimensionality of the rest of the brain to 100 components using principal components analysis (PCA). Next, we computed the Pearson correlation distance between every voxel in the MFC mask with each whole-brain PCA component,” – lines 121-127*

*“Should provide data or DB version #; should clarify whether you were using the public side or the core tools”*

We have noted the version number of the Neurosynth database and denoted that we are using the core python tools:

*“We analyzed version 0.4 of the the Neurosynth database…” - line 101*

*“Analyses were performed using the core Neurosynth python tools” – line 108*

*“Given the comment about open sharing, i was surprised that there was no mention of sharing these maps (or the code) w the community, as that would massively enhance the ultimate significance and value of these analyses.”*

We have clarified that we will share both the images and analyses by providing a more specific URL:

*“… code to replicate these analyses on any given region of interest are available as IPython Notebooks as part of the Neurosynth code base (https://github.com/neurosynth/neurosynth-mfc). “ – lines 108-109*

*“How were anatomical locations determined, e.g., via an automated labeling algorithm (AAL), standardized coordinate database (Talairach daemon), probabilistic atlases, etc.?*”

*“is it really appropriate to label the yellow and green regions in the right panel as 'pre-SMA”  
“*given Vogt's work cited in the intro, is 'pgACC' more appropriate than 'rACC'” *“Vogt has strongly argued for dropping the term dACC”*

We have more clearly detailed our method for labeling clusters. We have taken the reviewer’s suggestion to use the Harvard-Oxford anatomical atlas to more precisely localize regions. Alphanumeric labels were given to regions in the nine-clusters solutions to further avoid ambiguity and subjectivity. We clearly describe our procedure in pp. 9 in the Methods section:

*“To understand the anatomical correspondence of the resulting clusters, we calculated the probability of voxels in each cluster of occurring in probabilistic regions from the Harvard-Oxford atlas (H-O). We refer to H-O’s Juxapositional Lobule Cortex as Supplementary Motor Area (SMA) for consistency. We also compared the location of clusters to regions from cytoarchitechtonic atlases of medial motor areas (Picard and Strick, 1996), mid-cingulate cortex (Vogt, 2009) and vmPFC (Mackey and Petrides, 2014). To be precise, sub-regions in the nine-cluster solution were given alphanumeric names in addition to subjectively defined names.“ – line 152-18*

This approach revealed that all sub-regions in the middle zone were indeed probabilistically assigned primarily to the cingulate or paracingulate sulci. As such, we have renamed these clusters. Moreover, more careful comparison with Picard & Strick (1996) suggested “SMAr” should be re-named to “pre-SMA”. Finally, we have renamed “rACC” to “pgACC”, in concordance with the reviewer’s suggestion based on Vogt (2008).

Relevant updated section from the Results:

“*Within the posterior zone, we identified two clusters (Figure 2A; SMA [P1] & pre-SMA[P2]) with a high probability of occurring in SMA according to H-O. The two clusters delineated from each other near the vertical commissure anterior (VCA), consistent with cytoarchitechtonic delineations (Picard and Strick, 1996). However, SMA [P1] spanned multiple cytoarchitechtonic areas, extending ventrally to include portions of Picard & Strick’s cingulate zones.*

*In the middle zone, we identified four clusters consistent with midcingulate cortex (MCC). In particular, two anterior and two posterior clusters delineated from each other a few millimeters anterior to the VCA, consistent with Vogt’s definition of anterior and posterior midcingulate cortex (Vogt, 2009). The two dorsal clusters (pdMCC [M1] & adMCC [M2]) showed a high probability of occurring in H-O’s paracingulate gyrus, whereas the two ventral clusters (pvMCC [M3] & avMCC [M4]) showed a high probability of occurring the cingulate gyrus proper. Unlike some cytoarchitechtonic studies, we did not identify any regions exclusively were located in the cingulate sulcus, such as the rostral cingulate zone.”* – pp. 15-16

The reviewer also pointed out that we should clearly state the standardized stereotaxic space we are using. We now discuss this in more detail on line 105

*“As peak activations are populated automatically, the database does not differentiate between Talairach and MNI coordinates; however, all activations are treated as MNI coordinates, and all of the following analyses are in MNI152 coordinate space.”*

***“****a second concern is that the authors are in danger of interpreting the null, because (if I understand things correctly) they did not perform contrasts that would license the statements in the discussion about 'more' and 'greater'. “*

We thank the reviewer for catching this oversight, and have addressed this concern by calculating 95% confidence intervals for the log odds-ratio loading between topics and regions. Thus, if the 95% CI of log odd-ratio for a specific topic is non-overlapping between two regions, we can tentatively argue that the association strength between that topic and the two regions differs. In order to reduce the number of post-hoc exploratory comparisons made, we only report 95% CIs when we discuss any comparative relationships between regions and function with respect to the 11 topics present in Figure 4 (which were not selected on the basis of relatively differences between regions). ~~For example, we report 95% CIs for working memory, pain, reward and fear for all of the sub-regions in the middle zone, as the results from Figure 4 suggests potential differences in these topics between these sub-regions.~~

We wanted to be careful with these post-hoc exploratory tests, so we ensured to clearly label them as such in the results section:

*“We additionally report 95% confidence intervals of LORs whenever we comparatively discuss sets of regions. As these comparisons are post-hoc and exploratory, caution interpretation is warranted.* “ – lines 349-352

Relevant comparative passages from page 21:

*“However, exploratory post-hoc tests indicated SMA [P1] was more strongly associated with pain, while pre-SMA [P2] was more strongly associated with working memory (WM) (95% CI LOR. ‘pain’: SMA [0.6, 1.1], pre-SMA [-0.1, 0.4]; ‘WM’, SMA [-0.2, 0.1], pre-SMA [0.2, 0.4]).” – pp. 21*

*“However, post-hoc exploratory tests indicated dorsal MCC (M1 & M2) was more strongly associated with WM than ventral MCC (M3 & M4) (95% CI LOR. ‘pdMCC [0.5, 0.8], adMCC [0.4, 0.6], pvMCC [0, 0.15], avMCC [0, 0.3]) whereas ventral MCC showed a stronger association with affect than dorsal MCC (95% CI LOR. ‘fear’: pdMCC [-0.1, 0.4], adMCC [-0.4, 0.2], pvMCC [0.7, 1.2], avMCC [0.4, 0.9]; ‘reward: pdMCC [-0.4, 0.1], adMCC [-0.4, 0.1], pvMCC [0.3, 0.7], avMCC [0.3, 0.8]; ‘pain’: pdMCC [0.3, 0.8], adMCC [0.2, 0.7], pvMCC [1.1, 1.5], avMCC [0.6, 1.1]). Finally, both anterior clusters showed a greater association with decision-making than their posterior counterparts (95% CI LOR. pdMCC [-0.2, 0.3], adMCC [0.3, 0.8], pvMCC [-0.2, 0.4], avMCC [0.6, 1.1])”*

***“****In the anterior zone … activity across all three sub-regions was significantly predicted by episodic memory and social processing; however, the association with social processing was maximal for dmPFC [A3] (95% CI LOR. dmPFC [1.3, 1.7], pgACC [0.7, 1], vmPFC [0.6, 1]).”*

**Discussion**

*“a. The Discussion needs work. At times, the authors fall prey to over-selling the novelty and significance of their results****.*** *The results themselves are genuinely interesting and exciting, so there is no reason to exaggerate differences with prior research or to over-interpret the theoretical significance. “*

We have scaled back claims of theoretical novelty. For example, in the discussion of the middle zone, the reviewer convinced us that our results are in general consistent with the hypothesis that pain, negative affect and cognitive control consistent activate the MCC, and we now point that out clearly. However, we also more clearly outline novel implication of our results, such as the finding that reward was consistently associated with ventral MCC, suggesting this region may more generally integrate affect with cognitive control.

“*In contrast to claims of pain-specificity in MCC (Lieberman and Eisenberger, 2015), all four middle sub-regions were associated with pain and cognitive control. This finding is broadly consistent with adapative control hypothesis, which postulates that MCC integrates negative affective signals with cognitive control in order to optimize actions in the face of action-outcome uncertainty (Shackman et al., 2011; Cavanagh and Shackman, 2015). – lines 462-465*

*“Importantly, ventral MCC showed a preference not only for negative affect and pain, but also reward. Thus, the present results suggest that ventral aspects of MCC may incorporate low-level affective signals into cognitive control, whereas dorsal MCC may be more important for aspects of cognitive motor control that require working-memory or resolving interference.” – lines 469-473*

Following the reviewer’s suggestion, we have also moderated our claims regarding functional-anatomical specificity throughout the manuscript. For example, instead of claiming our results suggest additional “functional-anatomical specificity”, we instead say our results “demonstrate additional functional differences between sub-regions of MCC” (line 437).

In addition, we have heeded suggestion to more prominently discuss the finding that “no region is selectively activated by a single psychological concept” by expanding and moving that discussion to the second paragraph in the discussion (Line 648). It now reads:

*“Despite this progress, a number of important challenges remain for future research. Although the present analyses revealed distinct functional profiles for each region in MFC, it is notable that no region is selectively activated by a single psychological concept. This functional diversity is evident in the observation that at least two distinct topics significantly predict activation in each cluster and the relatively poor ability of our classifier’s to predict activation using only the single most strongly associated topic for each region. Our results suggest a complex many-to-many mapping between functionally separable brain regions and cognitive processes, in in contrast to recent claims of functional selectivity in MFC (Eisenberger, 2015; Lieberman and Eisenberger, 2015). Additional work is needed to determine the extent to which small ‘islands of specialization’ are indeed selective to specific processes, as has been argued using large-scale meta-analysis (Anderson et al., 2013; Yeo et al., 2014). “ – lines 471-482*

*“b. At times, the literature cited in the first half of the Discussion seemed dated”*

We have added updated the literature throughout the manuscript and now include a number of recent references, including papers by Shenhav, Botvinick & Cohen (2013); Mackey and Petrides (2014); Vogt (2009); Cavanagh and Shackman (2015); Eisenberger (2015); Cole et al. (2014); Mattar et al. (2015) and Hutchison et al. (2013).

*“c. The limitations section needs work. Careful scrutiny suggests that these are not necessarily the most important limitations and the most important avenues for future research.*

We have expanded and more clearly labeled the “future challenges” section (line 470). We have distinguished the limitations of Neurosynth from more general limitations of fMRI (lines 499-516). We now also discuss in more detail the possibility that our results suggest a complex many-to-many mapping between regions and functions and a dynamic interaction between brain organization and behavior.

Relevant section from the Discussion:

“*Despite this progress, a number of important challenges remain for future research. Although the present analyses revealed distinct functional profiles for each region in MFC, it is notable that no region is selectively activated by a single psychological concept. This functional diversity is evident in the observation that at least two distinct topics significantly predict activation in each cluster and the poor ability of our classifier’s to predict activation using only the single most strongly associated topic for each region. Our results suggest a complex many-to-many mapping between brain regions and cognitive processes, in in contrast to recent claims of functional selectivity in MFC (Eisenberger, 2015; Lieberman and Eisenberger, 2015). However, it has been previously suggested that functional specificity increases as with smaller regions of interest (Poldrack, 2006); as such additional researcher with greater spatial resolution is needed to determine if there indeed small ‘islands of specialization’ exist (Anderson et al., 2013; Yeo et al., 2014).*

*Although the present results provide a comprehensive snapshot MFC function, many have argued that brain regions dynamically assume different roles (Shackman et al., 2015) and modulate their connectivity as a function of task demands (Cole et al., 2014; Mattar et al., 2015). Moreover, MCC is likely to be among the most heterogeneous brain regions (Anderson et al., 2013) as evidenced by its very high activation rate (Nelson et al., 2010a; Yarkoni et al., 2011). Thus, it is possible that the functional profiles presented herein represented an ‘average’; as such, regions may deviate from this average depending on task-demands. For example, it’s possible that ventral MCC co-activates more strongly with the amygdala during ‘fear’, but co-activates with posterior insula during ‘pain’. An interesting avenue of future research will be to precisely characterize how co-activation and functional patterns of MCC change as a function of context through large-scale meta-analysis and dynamic resting-state functional connectivity (Hutchison et al., 2013). ”* – lines 471-489.

*d. This study is significant, but the theoretical and translational implications of the work are not clearly outlined in the Discussion. “*

In addition to the changes mentioned above, we now suggest that the hypotheses from this study could be tested by 1) the development of novel fMRI studies from the hypotheses proposed by this study and 2) large-scale functional mapping to individual subject anatomy.

“*The present report also provides several hypotheses that can be more carefully tested in future studies using the candidate psychological states presented here. For example, we observed that ventral MCC showed a greater preference for affect than dorsal MCC. However, it is difficult to localize these processes to precise anatomical areas given the wide inter-subject variability in paracingulate anatomy (Paus, 2001). Thus, future studies could more precisely localize the neural underpinnings of these processes in a single sample with subject-level anatomical registration. Wide spread data sharing should also enable large-scale meta-analyses that combine functional and anatomical data to more precisely localize activity to fine-grained anatomy. “* – lines XX-XX

Moreover, we have attempted to more carefully outline the theoretical implications of our work throughout the discussion. For example, in line 442: “*Thus, the present results suggest that ventral aspects of MCC may incorporate low-level affective signals into cognitive control, whereas dorsal MCC may be more important for aspects of cognitive motor control that require working-memory or resolving interference.*”

**Figures**

The reviewer suggested we display a coronal slice in Figure 1 to better display our ROI. While we have retained the sagittal slice in Figure 1 (mainly due to space considerations—it is not possible to show the entire MFC in a single coronal slice), we have added coronal slices to Figure 2. This accomplishes the same goal while also displaying in more detail the anatomical extent of our clusters.

Furthermore, in response to a comment by the reviewer, we have also added the silhouette plot to Figure 2 as the last panel, rather than having two separate figures.

Moreover, the reviewer suggested we try to better match colors between the brain and polar plots in Figure 4. We have re-rendered the brain plots in all figures to more accurately portray the color scheme used throughout the manuscript.

**Other specific comments**

*"Since most researchers tend to be intimately familiar with one particular domain of cognition"*

*“i strongly agree, but would object to calling it 'cognition' ... maybe 'psychological domain, such as pain'”*

We now use the term “psychological domain’. We have also avoiding using the term ‘cognition’ throughout the manuscript and instead use ‘psychological states’.

*"To determine which voxels across the brain co-activated with each MFC parcel, we performed a meta-analysis resulting in whole-brain maps that indicate which voxels across the brain are active in the studies that activated each parcel."*

*“i'm confused; how is this different than the meta-analytic co-activation on page 7?*

*when i 1st read this my comment was -- 'this is not really a meta per se, it seems more like a 'contrast' or a 'meta-analytic contrast' (like a moderator analysis in classic meta)'*

*but then i went and studied figure 1 and realized that (i think; could be wrong) that you are actually describing two steps at once, a meta and a meta contrast; you need to clarify this for the reader”*

These are indeed meta-analytic contrasts to determine whole-brain differences between studies that co-activate with one region (e.g. posterior MFC) versus control regions (e.g. middle and anterior MFC). We conducted these meta-analytic contrasts in order to highlight the differences between sets of related clusters. Thus, in the 3-cluster solution, we contrast the co-activation patterns of the three clusters with one other, whereas in the 9-cluster solution, we contrast the co-activation of clusters that correspond to the same zone to each other (e.g. vmPFC vs dmPFC & pgACC). We have tried to more clearly explain our methods in lines 231-244, and the caption for Figure 3: *“Meta-analytic co-activation contrasts for (A) three zones and B) nine sub-regions. Colored voxels indicate significantly greater co-activation with the seed region of the same color (at right) than control regions in the same row.”* We have also avoided using the term “unique”.

*“here you insert the additional adj 'specialization,' but given recent critical conversations in the blogosphere, might be better to either drop or use 'func preference profiles'”*

We have removed the term “specialization” for the manuscript and now use “functional preference profiles” instead.

**Reviewer 2**

*“The authors stress repeatedly the alignment of their findings with previous anatomical MFC studies "to a very substantial degree". Could they be more specific and provide evidence for this assertion. Does the number of clusters align with previous findings? However then I would expect them to find e.g., three distinct cingulate motor areas (as for example Dum & Strick). Or do the authors think that the spatial extent and location of their sub-areas resonates with previous research? Would they be able to demonstrate this? Or does their functional specialization analysis align with previous neurophysiological studies?”*

We have taken multiple steps to address this concern. First, we have more carefully outlined the extent to which our parcellation agrees with previous organizational schemes of MFC (see lines 423-449). In general, we find instances where our parcellation is quite similar to cytoarchtechtonic and connectivity-based approaches, such as the division between SMA and pre-SMA near the VCA. However, we also find several instances of disagreement. For example, as the reviewer notes, we did not identify three distinct cingulate motor areas. In fact, our most posterior cluster spans both SMA and the caudal cingulate zone. As such, we have tempered claims of substantial alignment between the present parcellation and previous studies. For example:

*“Although we did not expect our clusters to necessarily conform to morphologically derived regions, we nonetheless observed moderate correspondence. Within the posterior zone, we identified two clusters (Figure 2A; SMA [P1] & pre-SMA[P2]) with a high probability of occurring in SMA according to H-O. The two clusters delineated from each other near the vertical commissure anterior (VCA), consistent with cytoarchitechtonic delineations (Picard and Strick, 1996). However, SMA [P1] spanned multiple cytoarchitechtonic areas, extending ventrally to include portions of Picard & Strick’s cingulate zones.” –lines 267-273*

In addition, we have more thoroughly attempted to motivate co-activation based parcellation in the introduction by noting the limitation of previous studies (lines 71-73). In particular, many previous studies indirectly infer functional differences from morphological or connectivity differences, but since they do not directly measure how the MFC responds to various challenges, they cannot directly determine if putative sub-regions are ‘functionally different’. A priori, there is no particular reason to expect very strong (e.g., one-to-one) mappings between anatomically or cytoarchitectonically-defined clusters and functionally-defined clusters. For instance, two parts of MFC that contain neurons with similar morphological distributions could potentially play very different roles in cognition in virtue of having different connectivity patterns with the rest of the brain. We believe co-activation based parcellation provides a more direct window into functional differences across different parts of MFC.

*“What does their method of using data-mining fMRI activation peaks to the above mentioned sizeable literature on MFC sub-specialisation? I presume that their method does not allow for finer-grained sub-divisions than cyto-architecture, receptor density or tracer injection based studies? If they wanted for a function-based subdivision could they not have used a "functional localizer" approach as Amiez and Petrides (2014)?”*

It is true that our analyses are limited in spatial specificity by the limitations of fMRI itself and of our meta-analytic data. However, we do not see this as a principled reason to abandon such an approach in favor of other methods. As noted above, we think it is unlikely that there is a single correct parcellation common to different methods of analysis. Our expectation is that a coactivation-based parcellation would inevitably produce somewhat different results from parcellations based on cytoarchitectonics, receptor density, gene expression, etc., no matter how fine-grained the data in question were. As we have clarified above, we do not see the goal of this parcellation (or any other) as being to arrive at *the* single true parcellation of the MFC, because we do not think such a thing exists. Rather, our effort is designed to help understand how different sectors of the MFC contribute functionally to different aspects of cognition and behavior.

What the large-scale meta-analyses conducted in the present work allow us to do is better understand the functional significance of the resulting clusters across a wide variety of psychological manipulations. While we think functional localizers are an excellent approach when researchers are focused on narrowly-defined aspects of cognition (e.g., face perception, motor responding, etc.), such localizers are necessarily constrained to only consider a small subset of possible psychological manipulations. For example, in the Amiez & Petrides (2014) study the reviewer cites, the authors exclusively used motor localizers (e.g., for the arm, hand, foot, etc.). We think that this is precisely the right approach if one’s goal is to understand how different cingulate regions contribute specifically to motor control; however, it does not provide insights into the large-scale fractionation of MFC in the context of domain-general cognition. Moreover, one unique benefit of using a database that spans a very broad range of functional tasks is that, unlike studies using functional localizers, we are able to tackle the ‘reverse inference’ problem by estimating the degree to which a region is *preferentially* recruited by a particular process. As we note in the introduction, this is particular problematic for areas with a high rate of activation across studies, like MCC / pre-SMA. Such regions are likely to activate in a wide range of localizer tasks, potentially leading researchers to conclude that they are selective for the particular localizers used, when in fact they show similar affinity for a wide range of other processes. We have attempted to make this point more clear in the introduction by unpacking this problem:

*“There have been several recent attempts to define distinct sub-regions of MFC by inferring functional differences on the basis of morphology (Vorobiev et al., 1998; Vogt, 2009; Palomero-Gallagher et al., 2013) and in-vivo structural (Beckmann et al., 2009; Sallet et al., 2013; Neubert et al., 2015) and functional connectivity (Andrews Hanna et al., 2010). Although such studies have greatly advanced our understanding of MFC’s organization, these methods cannot directly demonstrate that the resulting regions are functionally separable, nor reveal their functional significance as they do not measure their response to various psychological challenges.” – lines 51-58*

*“I am not sure if I follow the assertion: "Although the 12-cluster solution results in a marginally better silhouette score, this comes at the cost of additional complexity." Why would they discard this solution if it fits the criteria that they themselves set better? If they think that MFC organization is indeed more complex why would this be a cost?”*

The Reviewer points to a difficult general issue that faces virtually any parcellation effort: there are many different criteria for selecting a “good” parcellation, and it is rarely clear how to define a cost function that optimizes all of the relevant constraints. Our view is that individual metrics like the silhouette score (and there are a large number of such metrics one could use; cf. Craddock et al., 2012) should guide, but not deterministically dictate, decisions about parcellation schemes. One particularly common issue with such metrics is that they often are insensitive to human constraints on understanding (if our analysis had suggested an optimal *k* of 45, we would not want to present in our paper results for 45 different clusters!). Thus, we feel that there is nothing inherently wrong with combining quantitative metrics with subjective judgment in this context. In the previous version of the manuscript, we elected to focus on a 9-cluster solution rather than a 12-cluster solution because the improvement in silhouette score was negligible, and the increase in complexity was appreciable. This decision does not imply that we believe a 9-cluster solution to be “truer” than a 12-cluster solution; it is simply a recognition of the fact that there are multiple constraints on what constitutes a practically useful parcellation, and one of them is complexity. Had the silhouette score profile looked different (e.g., if the silhouette score had been much greater for the 12-cluster solution than a 9-cluster solution), we would probably have made a different choice.

That said, we agree with the reviewer that our reasoning for choosing the 9-cluster solution was not made sufficiently clear in the manuscript. We have therefore made two changes. First, we now clarify the motivation for choosing to focus on the 9-cluster solution in lines 250-255:

“The plateauing of silhouette scores suggests that there is little objective basis for selecting one solution over another past around 9 clusters (Thirion et al., 2014); we have therefore opted to focus on the 3-cluster and 9-cluster solutions because they provide greater theoretical parsimony than more fine-grained solutions at two spatial scales.” –lines 243-247

Second, to make sure that readers do not come away thinking that we cherry-picked the value of 9 for arbitrary reasons (rather than because it is the smallest number that shows a negligible difference in silhouette score), we now include the 12-cluster solution as part of Figure 2.

*“Are there contextual differences in co-activation patterns? E.g., dACC appears to co-activate with DLPFC and amygdala. It also appears to be associated with conflict, decision making and pain. Is it more activated with the amygdala in studies that mention pain and more activated with DLPFC in studies that mention conflict?”*

The reviewer raises an interesting question, and one that we have begun to explore in other contexts. The results presented here focus on providing an overall picture of the co-activation and function of MFC. We agree with the Reviewer that it is very likely that certain regions are dynamically involved with different processes, depending on the context. However, this question is out of the scope of the present report, and would require extensive further research (there are also technical complications in modeling context-specific coactivation in this case, due to the meta-analytic nature of the data). We have noted this as a potential avenue for future research in a new paragraph in the discussion:

*Although the present results provide a comprehensive snapshot MFC function, many have argued that brain regions dynamically assume different roles (Shackman et al., 2015) and modulate their connectivity as a function of task demands (Cole et al., 2014; Mattar et al., 2015). Moreover, MCC is likely to be among the most heterogeneous brain regions (Anderson et al., 2013) as evidenced by its very high activation rate (Nelson et al., 2010a; Yarkoni et al., 2011). Thus, it is possible that the functional profiles presented herein represented an ‘average’; as such, regions may deviate from this average depending on task-demands. For example, it’s possible that ventral MCC co-activates more strongly with the amygdala during ‘fear’, but co-activates with posterior insula during ‘pain’. An interesting avenue of future research will be to precisely characterize how co-activation and functional patterns of MCC change as a function of context through large-scale meta-analysis and dynamic resting-state functional connectivity (Hutchison et al., 2013). ” ­– lines 483-489*

*“Even though the authors state distinct areas with distinct functions and connections there appear to be strong overall similarities in neighboring regions in co-activation and function potentially with gradients of change along different axes. For example "motor" seems to gradually decrease from posterior to anterior. Pain appears to decrease from ventral to dorsal Similarly DLPFC co-activation appears to increase from posterior to anterior. It would be very interesting to see if there is concordance or correlation in these functional - connectivity gradients / changes? E.g., a gradient of decrease in pain is associated with a decrease in amygdala co-activation?”*

We have attempted to discuss potential gradients in more detail in the discussion. First, we have tried to formally identify functional gradients by introducing post-hoc exploratory tests of functional differences between sub-regions in the results (lines 381-400). These tests reveal some potential gradients, such as the association with ‘reward’ becoming stronger ventrally in the anterior zone. Although we do not have a test to formally determine if these gradients are accompanied by specific changes in co-activation, we have attempted to discuss such possibilities in the discussion:

* “In contrast, pre-SMA (P2) showed a stronger association with cognitive control and co-activated with regions important for goal-directed cognition (e.g. DLPFC, aIns).” Lines 427-428
* “Notably, both dorsal MCC clusters were more strongly associated with WM – and showed great co-activation with fronto-parietal control regions and aIns— while ventral MCC was more strongly associated with affect and co-activated more strongly with subcortical regions, such as amygdala and striatum.” Lines 438-440

*Why does the three-zone subdivision group together regions with vastly different cyto-architecture and separate regions with similar cyto-architecture?*

We think this concern is addressed by the more extensive discussion in our introduction (and above) about the limitations of cytoarchitechtonic-based parcellations. In short, we don’t believe that there is a reason to expect close agreement between these two methods, as differences in cytoarchitechtonic properties need not necessarily translate to functional differences at the level of cognition, or vice versa. Moreover, as the three-zone solution is so broad, it will necessarily group together regions that differ in morphological and cytoarchitechtonic properties.

*Minor:*

*How well did the Harvard Oxford grey matter match the implicit Neurosynth data-base grey matter? I suppose all Neurosynth foci should lie in the grey matter? What percentage of Neurosynth foci are outside the 30% Harvard-Oxford grey matter atlas?*

Because all Neurosynth data are masked by the MNI152 gray matter template bundled with FSL, there will necessarily be at least a reasonable correspondence with the Harvard-Oxford atlas. The Reviewer is correct that coordinates outside gray matter are deliberately excluded from the Neurosynth database. We now explicitly address this on lines 118-121: “*In general, Neurosynth’s activation mask corresponded highly with probabilistic locations of cerebral cortex with the exception of portions of precentral gyrus and far ventromedial prefrontal cortex– which showed low activation although they were more than 50% likely to be in cerebral cortex.*”