# Distinct neural contributions to confidence in detecting (but not discriminating) visual stimuli

# Pre-registered Analysis

##### Modulation of Confidence

The global confidence (GC) design matrix revealed linear modulation of confidence all across the brain for correct responses.

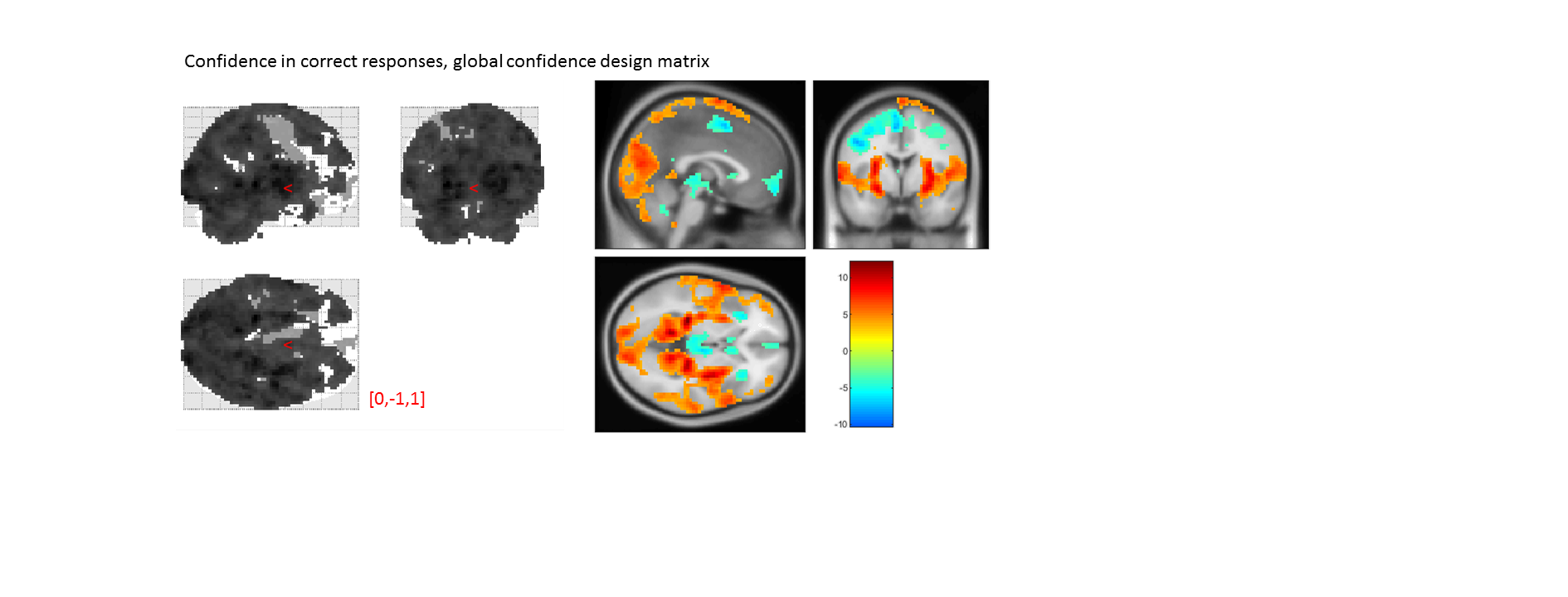


Figure S1: confidence modulation in correct responses, from the global-confidence design matrix. Uncorrected, thresholded at p<0.001.

From our pre-specified ROIs, only the vmPFC and BA46 ROIs showed a significant effect of confidence in correct responses, in the opposite direction to what we expected based on previous studies.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | average beta | T value | p value | Standard deviation |
| vmPFC | -0.34579 | -3.05815 | 0.004321 | 0.668941 |
| pMFC | -0.31008 | -2.48307 | 0.018119 | 0.738779 |
| precuneus | 0.247039 | 2.296966 | 0.027905 | 0.636276 |
| Ventral striatum | -0.05599 | -1.51007 | 0.140263 | 0.219349 |
| FPl | 0.164591 | 1.519002 | 0.138007 | 0.641035 |
| FPm | -0.11771 | -1.44594 | 0.157349 | 0.48162 |
| BA 46 | 0.366381 | 3.772601 | 0.000618 | 0.574548 |

##### Effect of task

No brain region presented a main effect of task in a whole-brain contrast of detection-discrimination. Similarly, none of our ROIs presented a main effect of task.

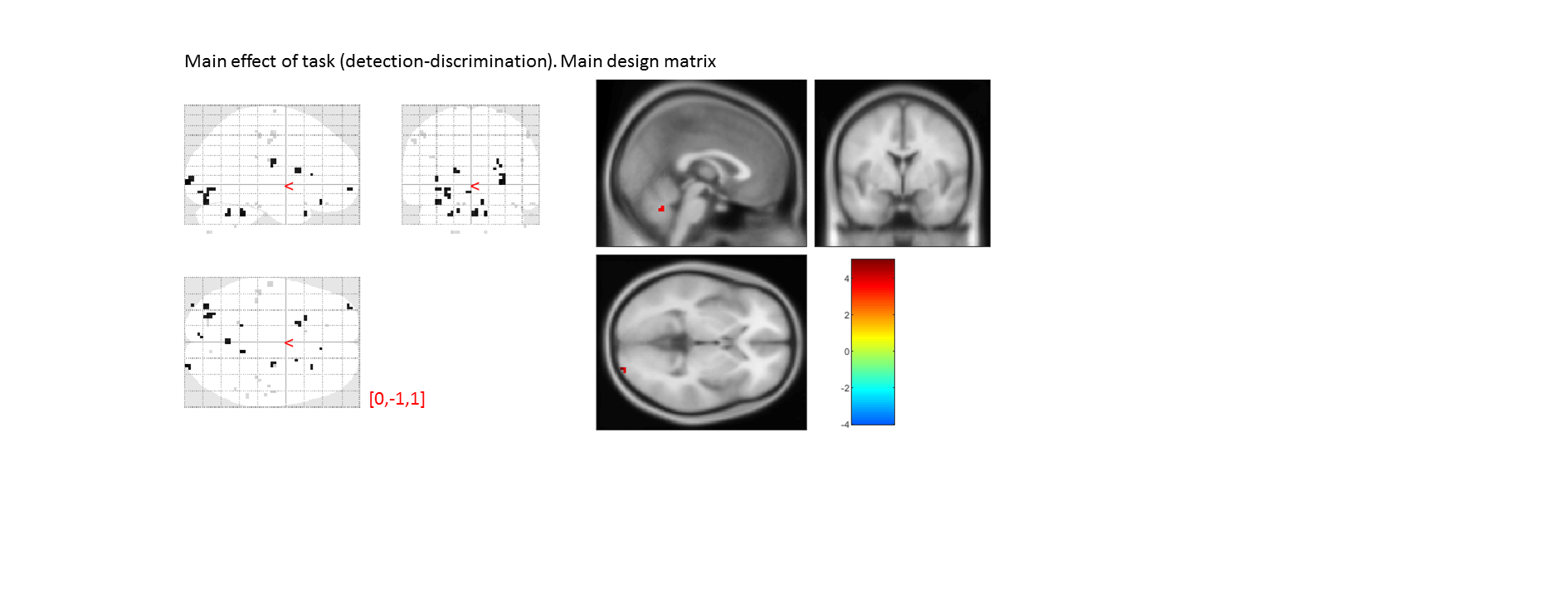


Figure S2: main effect of task, from the main design matrix. Uncorrected, thresholded at p<0.001.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | average beta | T value | p value | Standard deviation |
| vmPFC | -0.01315 | -0.04755 | 0.962352 | 1.635718 |
| pMFC | 0.147204 | 0.601448 | 0.551531 | 1.447956 |
| precuneus | -0.04469 | -0.16065 | 0.873323 | 1.64565 |
| Ventral striatum | 0.093292 | 0.769216 | 0.447076 | 0.717512 |
| FPl | 0.282758 | 1.077425 | 0.288876 | 1.55261 |
| FPm | 0.005014 | 0.024215 | 0.980823 | 1.224891 |
| BA 46 | 0.379071 | 1.18936 | 0.242538 | 1.885565 |

##### No-Yes responses

No brain region presented a main effect of task in a whole-brain contrast of no-yes responses, within detection. Similarly, none of our ROIs presented a main effect of response.

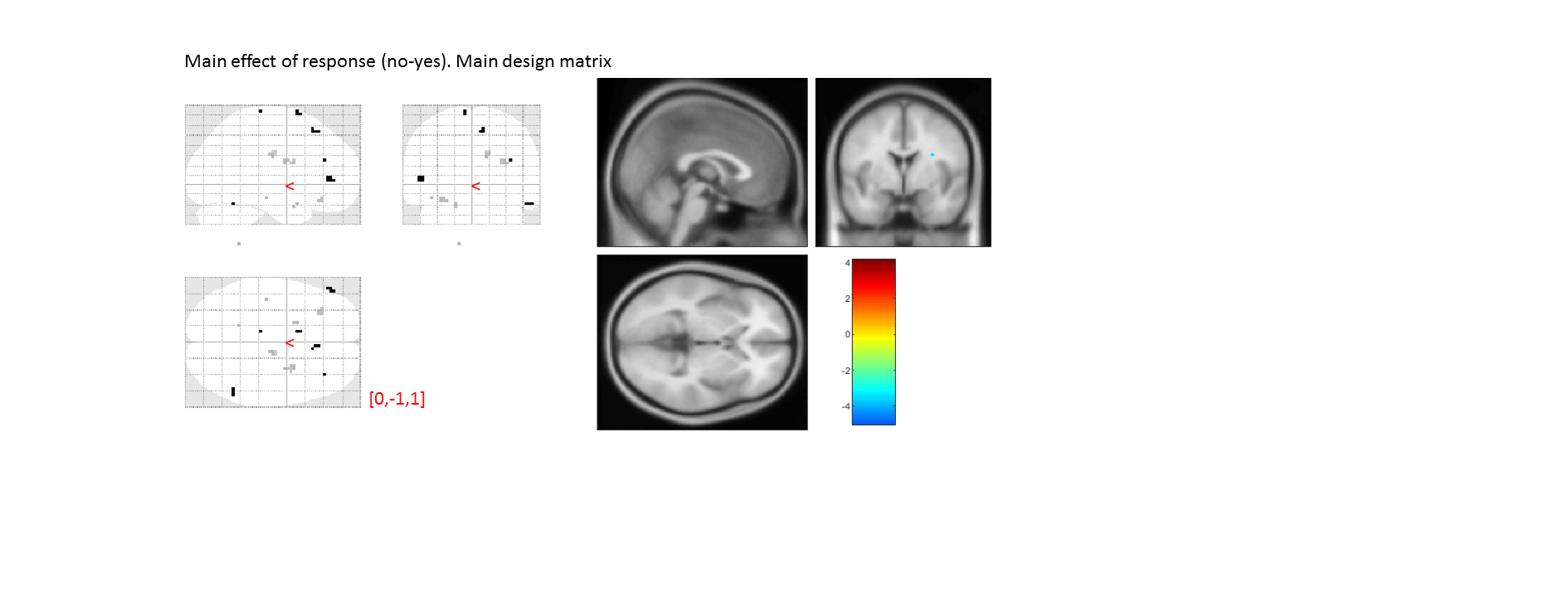


Figure S3: main effect of response, from the main design matrix. Uncorrected, thresholded at p<0.001.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | average beta | T value | p value | Standard deviation |
| vmPFC | -0.04999 | -0.31387 | 0.755538 | 0.942327 |
| pMFC | 0.137708 | 0.893458 | 0.377893 | 0.91184 |
| precuneus | -0.04233 | -0.52038 | 0.60617 | 0.481275 |
| Ventral striatum | -0.06841 | -1.75252 | 0.0887 | 0.23092 |
| FPl | 0.113597 | 0.902732 | 0.373019 | 0.744463 |
| FPm | 0.059862 | 0.510763 | 0.612818 | 0.693377 |
| BA 46 | 0.149013 | 1.228072 | 0.227854 | 0.717851 |

##### Confidence

We found the expected pattern of a positive linear modulation of confidence in the vmPFC and precuneus, and a negative modulation in the pMFC and lateral surfaces.

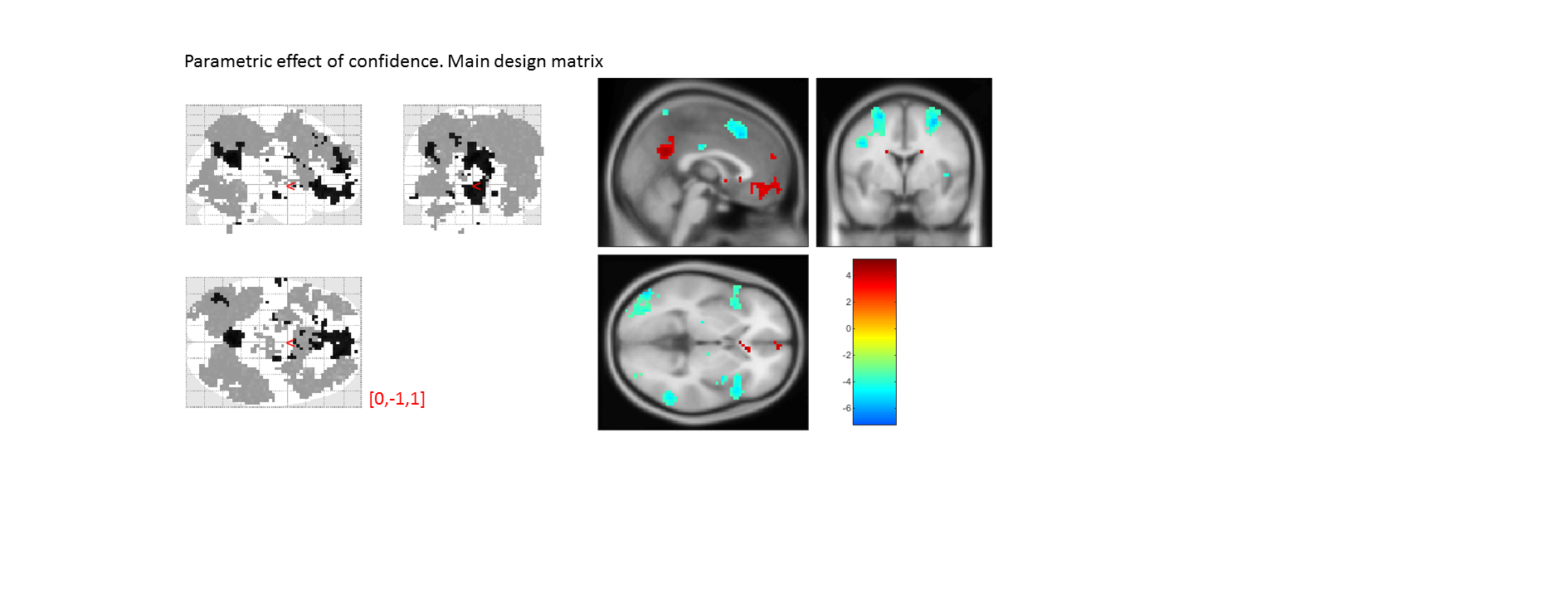


Figure S4: parametric effect of confidence, from the main design matrix. Uncorrected, thresholded at p<0.001.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | average beta | T value | p value | Standard deviation |
| vmPFC | 0.608913 | 3.207465 | 0.002916 | 1.123123 |
| pMFC | -0.72642 | -5.70501 | 2.08E-06 | 0.753294 |
| precuneus | 0.375672 | 3.058718 | 0.004314 | 0.726613 |
| Ventral striatum | 0.095389 | 1.726673 | 0.093303 | 0.326829 |
| FPl | -0.13936 | -1.03414 | 0.308373 | 0.797235 |
| FPm | 0.295691 | 2.496942 | 0.017533 | 0.70059 |
| BA 46 | -0.35397 | -3.62734 | 0.000929 | 0.577321 |

##### Metacognitive adequacy

Following Wokke, Cleeremans and Ridderinkhof (2017), we quantified metacognitive adequacy as the interaction between accuracy and confidence, such that metacognitive adequacy scales positively with confidence in correct responses, but negatively in incorrect responses. No brain region survived whole-brain correction for this contrast. Within our ROIs, we found a significant negative effect of metacognitive adequacy in the dorsolateral prefrontal cortex (BA46).

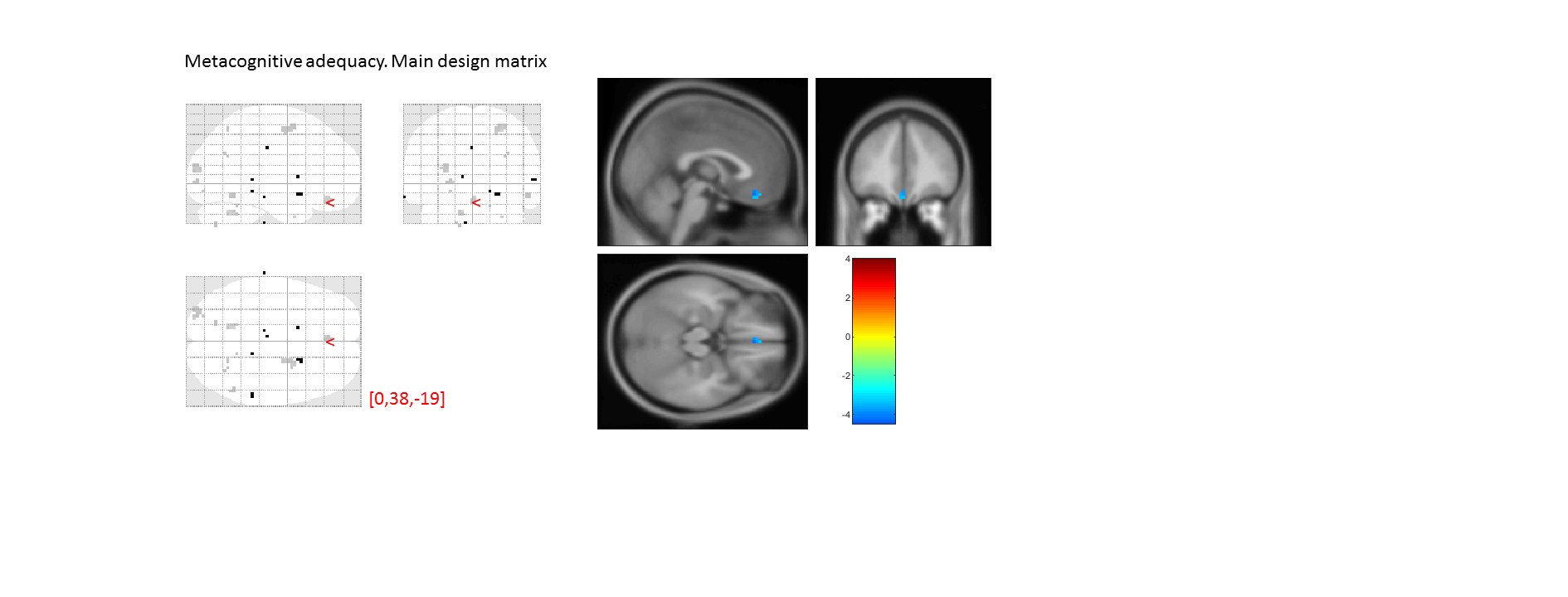


Figure S5: parametric effect of metacognitive adequacy, from the main design matrix. Uncorrected, thresholded at p<0.001.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | average beta | T value | p value | Standard deviation |
| vmPFC | -0.1031 | -0.91051 | 0.368962 | 0.669872 |
| pMFC | -0.0914 | -0.82244 | 0.416558 | 0.65746 |
| precuneus | -0.11388 | -1.06825 | 0.292932 | 0.630654 |
| Ventral striatum | 0.03709 | 0.941485 | 0.353098 | 0.233063 |
| FPl | -0.04155 | -0.29735 | 0.768012 | 0.826647 |
| FPm | -0.17904 | -1.15195 | 0.257379 | 0.919508 |
| BA 46 | -0.35397 | -3.62734 | 0.000929 | 0.577321 |

##### Confidence x task interaction

Contrasting the parametric modualtion of confidence between the two tasks revealed many small clusters that presented a more positive slope for confidence in detection, compared to discrimination, including in the right BA10 and the calcarine sulcus in the occipital lobe. None of these clusters reached significance at the whole-brain level.

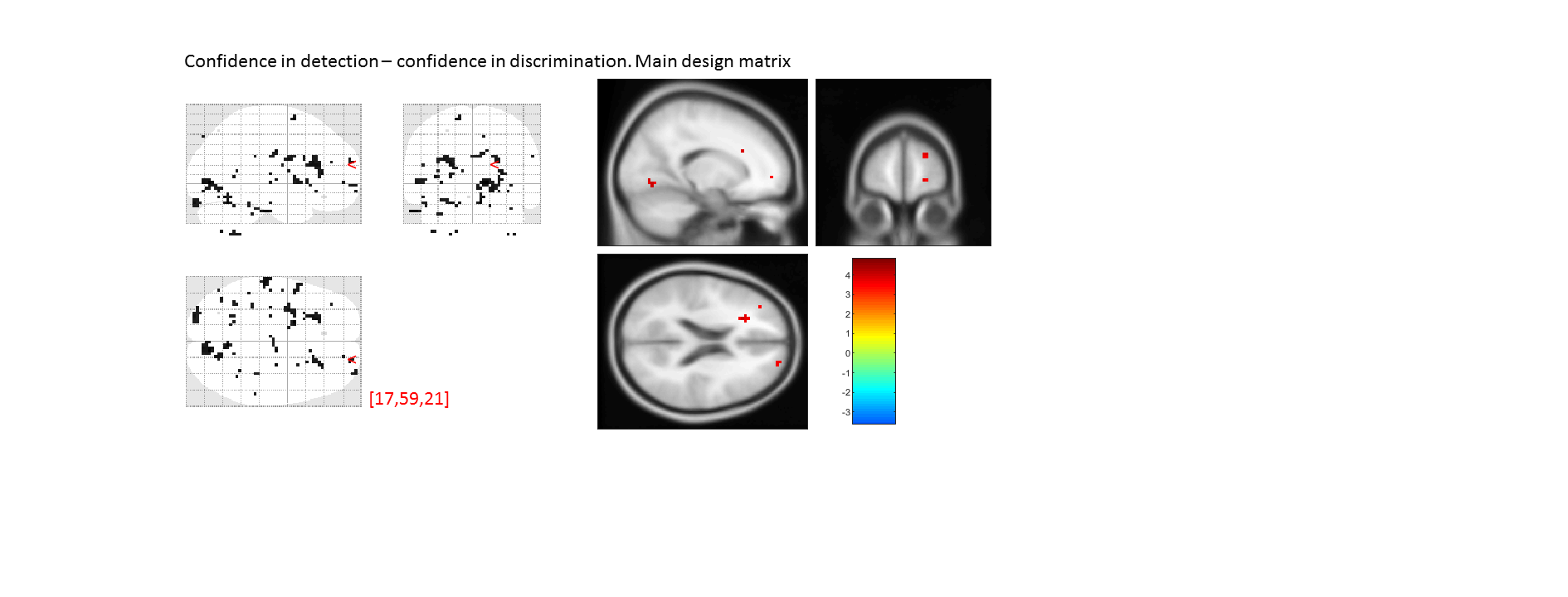


Figure S6: interaction between task and confidence, from the main design matrix. Uncorrected, thresholded at p<0.001.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | average beta | T value | p value | Standard deviation |
| vmPFC | -0.07672 | -0.51791 | 0.607877 | 0.876322 |
| pMFC | 0.025989 | 0.222144 | 0.825531 | 0.692131 |
| precuneus | -0.07994 | -0.71101 | 0.481926 | 0.665182 |
| Ventral striatum | 0.043207 | 1.041387 | 0.305045 | 0.245455 |
| FPl | 0.068901 | 0.739352 | 0.464768 | 0.551325 |
| FPm | -0.09755 | -0.84417 | 0.404474 | 0.683629 |
| BA 46 | 0.09987 | 0.842872 | 0.40519 | 0.700981 |

##### Metacognitive adequacy x task interaction

We found no significant effect of an interaction between metacognitive adequacy and task anywhere in the brain.

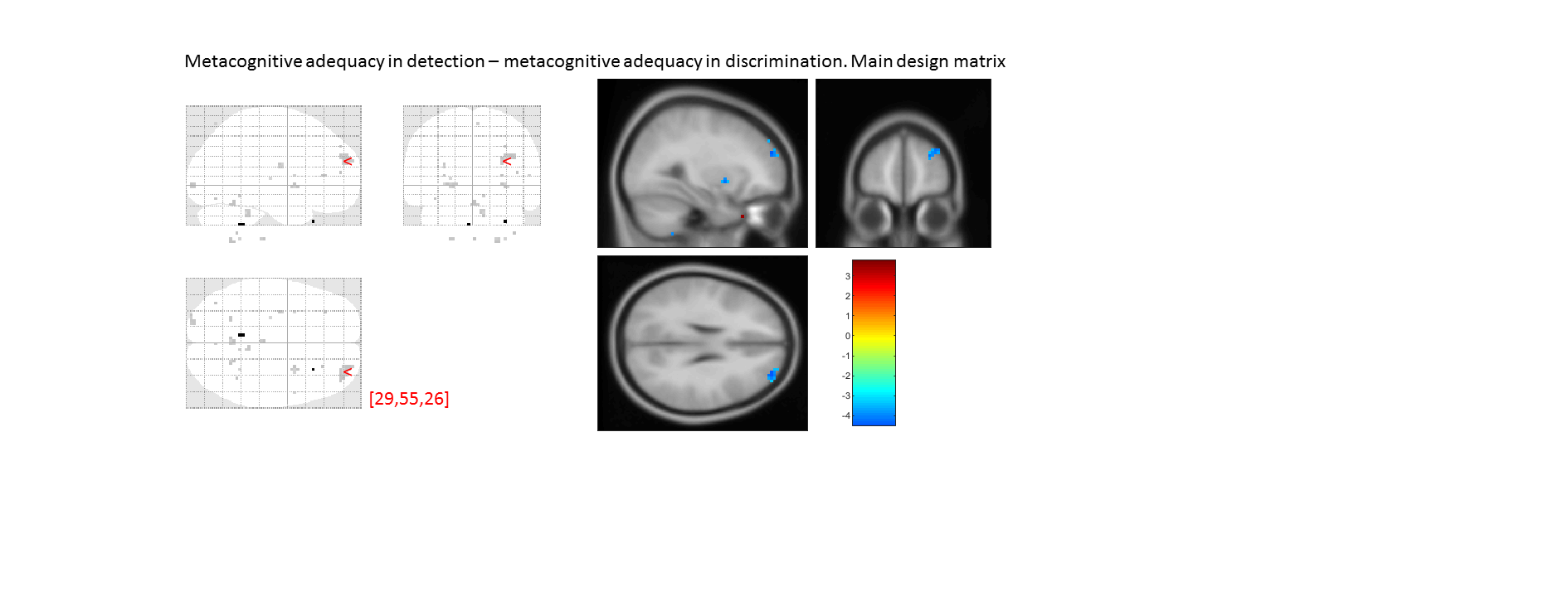


Figure S7: interaction between task and metacognitive adequacy from the main design matrix. Uncorrected, thresholded at p<0.001.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | average beta | T value | p value | Standard deviation |
| vmPFC | 0.09064 | 0.841626 | 0.405878 | 0.637138 |
| pMFC | 0.009987 | 0.077203 | 0.938915 | 0.765339 |
| precuneus | 0.189573 | 2.123296 | 0.041088 | 0.528202 |
| Ventral striatum | 0.038387 | 0.941033 | 0.353326 | 0.24133 |
| FPl | -0.01766 | -0.2368 | 0.814232 | 0.441082 |
| FPm | 0.069147 | 0.901441 | 0.373695 | 0.453805 |
| BA 46 | -0.1008 | -1.04149 | 0.304997 | 0.57261 |

##### Confidence x response interaction

A contrast of confidence in ‘no’ versus ‘yes’ responses revealed a more negative slope for ‘no’ responses in the right temporoparietal junction. Subthreshold clusters included the right and left superior temporal sulci, and the left frontal operculum (Broca’s area). None of our pre-specified ROIs was significant for this contrast, although notably all ROIs presented a negative effect, and the three subdivisions of the frontopolar cortex presented were all marginally significant for this effect.

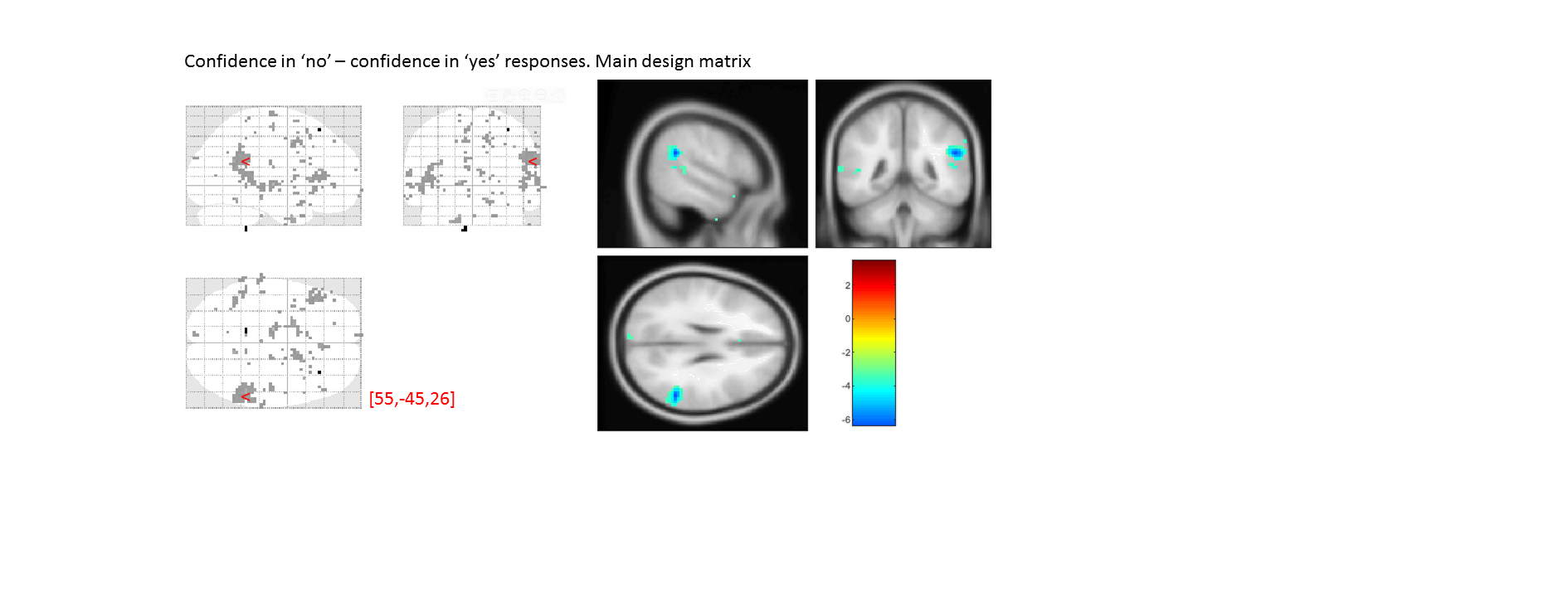


Figure S8: interaction between response in detection and confidence, from the main design matrix. Uncorrected, thresholded at p<0.001.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | average beta | T value | p value | Standard deviation |
| vmPFC | -0.11941 | -1.22917 | 0.227449 | 0.574733 |
| pMFC | -0.12918 | -1.76193 | 0.087071 | 0.433753 |
| precuneus | -0.052 | -0.71731 | 0.47808 | 0.428892 |
| Ventral striatum | -0.04356 | -1.54473 | 0.131668 | 0.166844 |
| FPl | -0.17959 | -2.3657 | 0.023839 | 0.449109 |
| FPm | -0.21068 | -2.02782 | 0.050473 | 0.614662 |
| BA 46 | -0.14502 | -2.00678 | 0.052778 | 0.427539 |

##### Metacognitive adequacy x response interaction

A contrast of metacognitive adequacy in ‘no’ versus ‘yes’ responses revealed no activations anywhere in the brain.

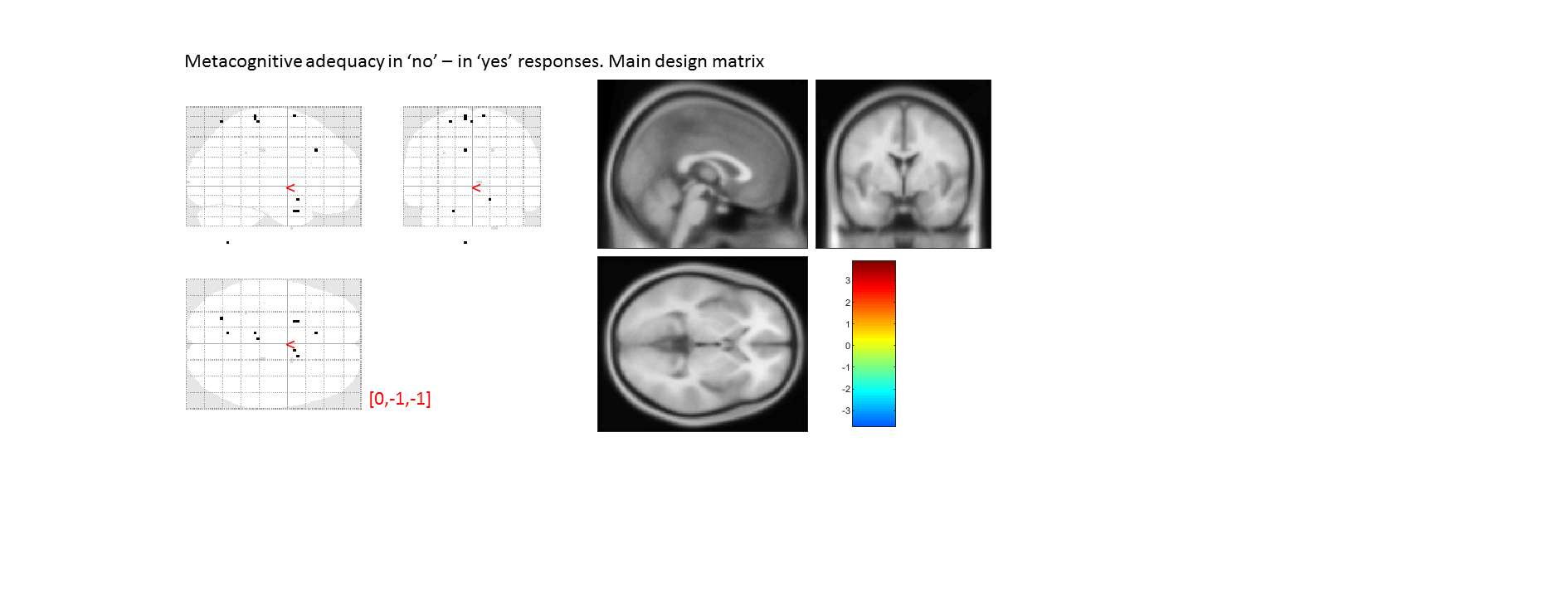


Figure S8: interaction between response in detection and confidence, from the main design matrix. Uncorrected, thresholded at p<0.001.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | average beta | T value | p value | Standard deviation |
| vmPFC | -0.0362 | -0.3132 | 0.756044 | 0.683821 |
| pMFC | 0.069815 | 1.1393 | 0.262542 | 0.362531 |
| precuneus | -0.10559 | -0.78435 | 0.438267 | 0.796435 |
| Ventral striatum | -0.02853 | -0.77488 | 0.443764 | 0.217847 |
| FPl | -0.0304 | -0.54826 | 0.587096 | 0.32808 |
| FPm | 0.020121 | 0.282576 | 0.779215 | 0.421248 |
| BA 46 | -0.0186 | -0.39863 | 0.692656 | 0.276084 |

#### Intersubject correlations

For each participant, we extracted task-specific M-ratios by dividing meta-d’ by d’ separately for each task. We then included these estimates as covariates into a group-level contrast of confidence in correct responses in detection and in discrimination. We found no significant correlations between task-specific M-ratio estimates and modulation of confidence anywhere in the brain.

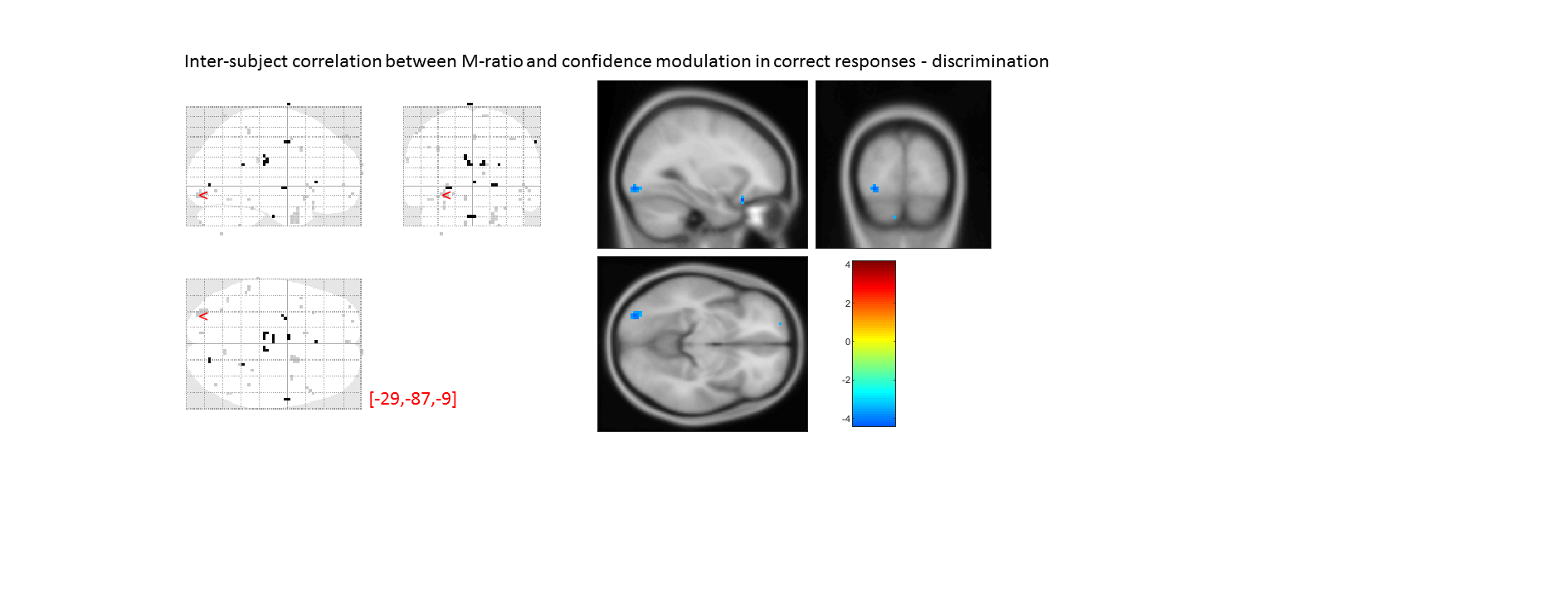


Figure S9: Parametric effect of M-ratio on the modulation of confidence in correct responses, in discrimination. Uncorrected, thresholded at p<0.001.

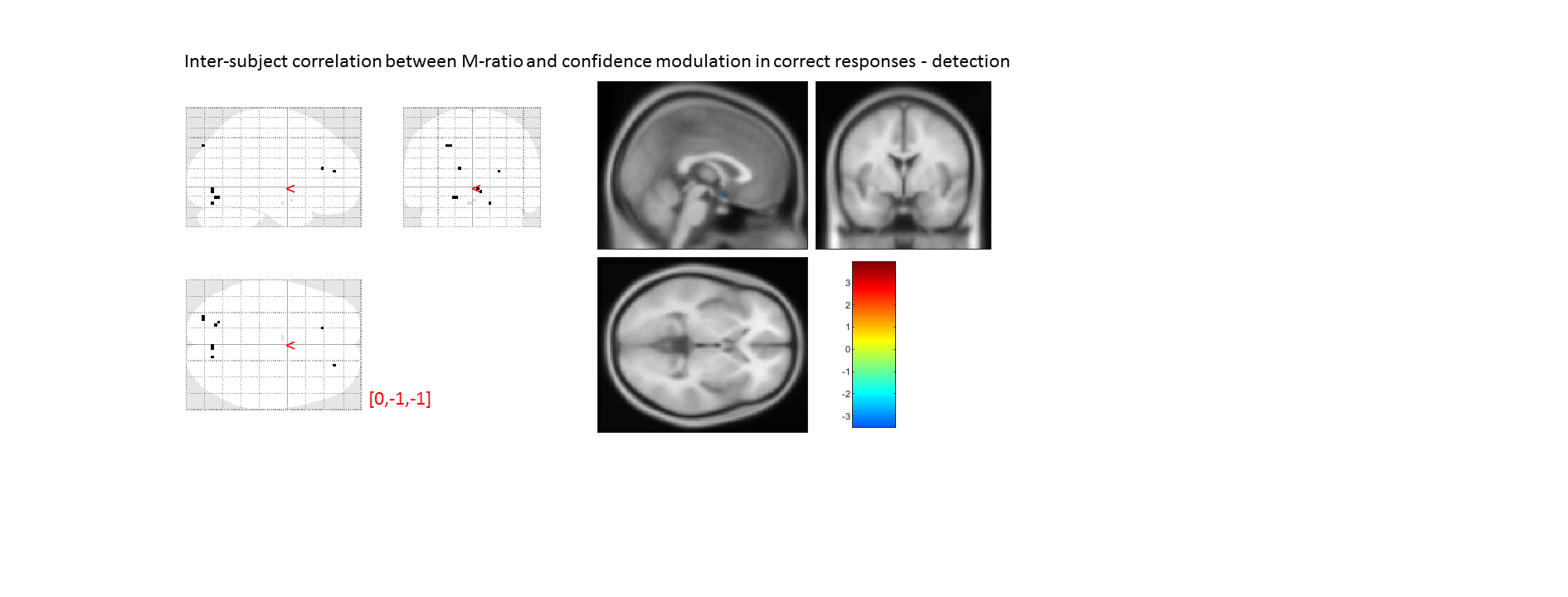


Figure S10: Parametric effect of M-ratio on the modulation of confidence in correct responses, in detection. Uncorrected, thresholded at p<0.001.

|  |  |  |
| --- | --- | --- |
|  | T value: Detection | T value: Discrimination |
| vmPFC | 0.9220 | -0.1053 |
| pMFC | -0.1114 | -0.5937 |
| precuneus | -1.1142 | 1.0797 |
| Ventral striatum | 0.4724 | -0.2499 |
| FPl | 0.4582 | -0.8919 |
| FPm | 0.8350 | -0.3248 |
| BA 46 | 0.3076 | -0.4302 |

##### 

#### Multivariate analysis

We used The Decoding Toolbox (Hebart, Gargen, & Haynes, 2015) for our multivariate analysis. We wish to draw the reader’s attention to the fact that while the mapping between confidence rating and visual feedback was counterbalanced between participants, it was not counterbalanced within participant. For this reason, shared multivariate representations of confidence between task and responses can theoretically reflect representations of the visual interface instead of some internal variable or process. Although this is unlikely for higher-order brain regions such as the frontopolar cortex, we decided to leave this analysis out of the main manuscript, and only include it here for completeness. For this analysis, we only used the data of 29 participants with four or five usable runs.

##### General confidence representation

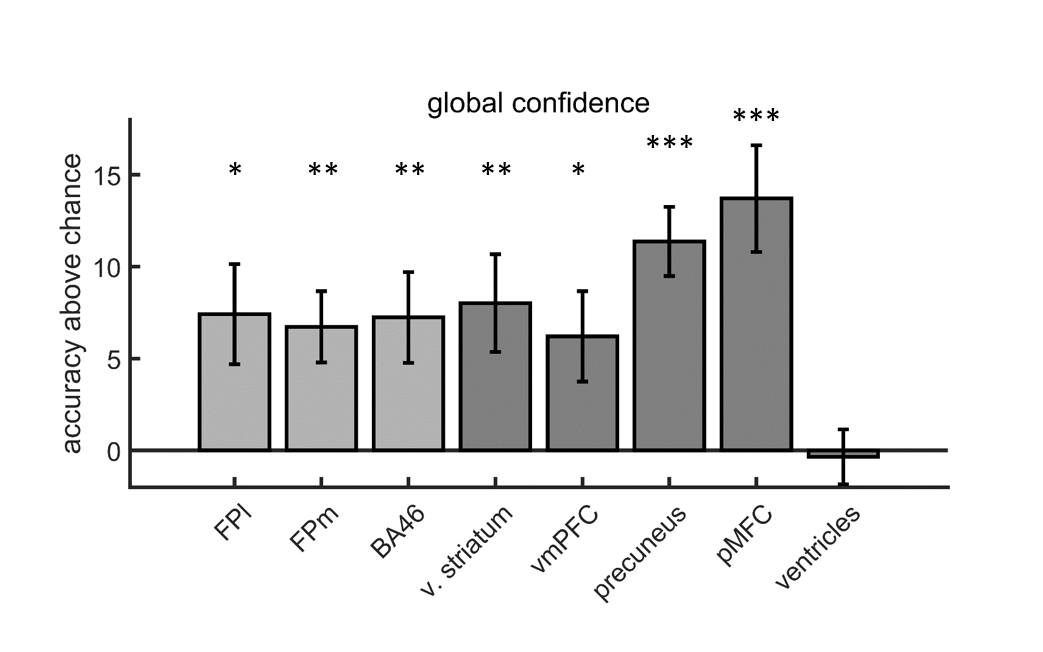


Figure S11: accuracy minus chance (50%) for classification of confidence (high vs. low; based on median split) in correct responses.

##### Task-specific and task-invariant confidence representation

We found no clear evidence for task specificity in the representation of confidence in any of our prespecified brain regions. In the vmPFC ROI, classification accuracy was higher for discrimination than for detection (t(28)=2.50, p<0.05), but this effect did not survive correction for multiple comparisons. We note that this effect aligns well with the univariate pattern observed in the vmPFC. The ventral striatum presented lower classification accuracy when trained and tested on different tasks (cross-classification), hinting to task-specificity in this region. However, this effect was not strong enough to survive correction for multiple comparisons (t(28)=2.34,p<0.05).

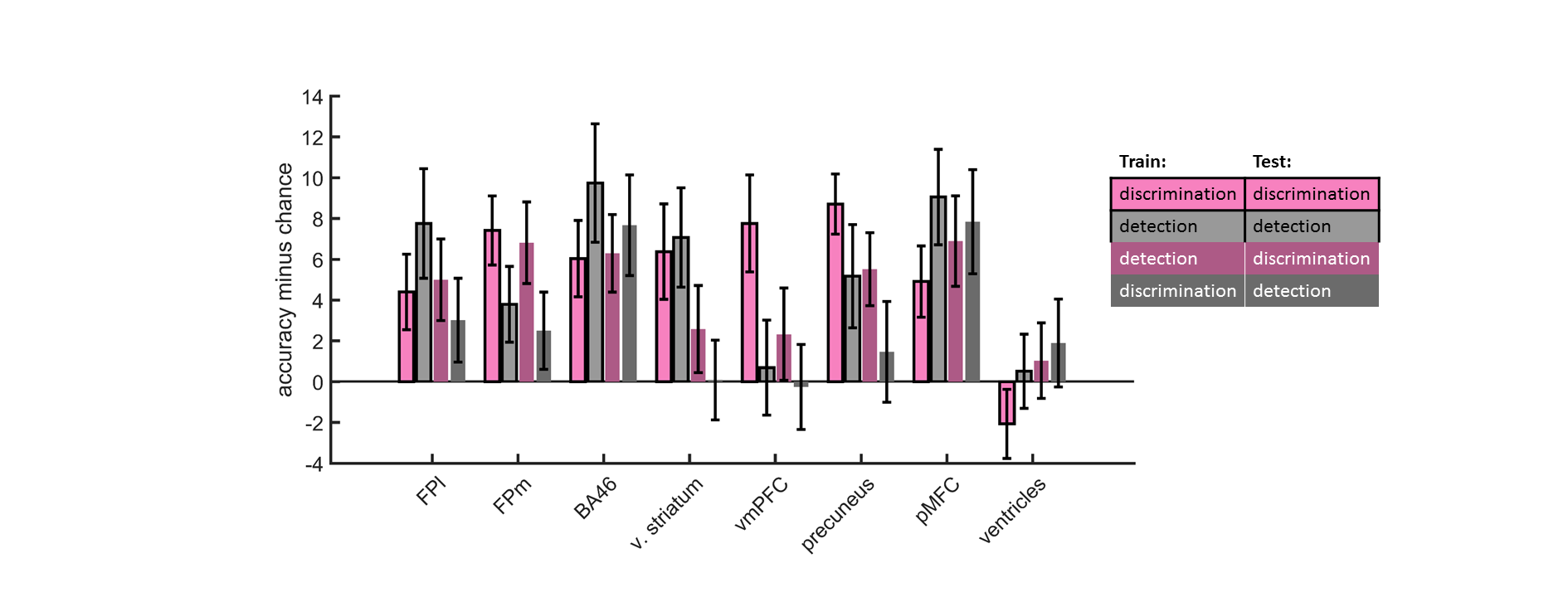


Figure S12: accuracy minus chance for classification of confidence (high vs. low; based on median split) in correct responses in discrimination (pink, black edges), and in detection (gray, black edges), and from a cross-classification between tasks: detection to discrimination (dark pink, no edges), and discrimination to detection (dark gray, no edges).

##### Shared representations of discrimination confidence and detection response

The purpose of this analysis was to identify representations of stimulus visibility. These representations should contribute to decision in detection where high visibility will be associated with ‘yes’ responses, and to confidence in discrimination where high visibility will be associated with high confidence ratings. Classification accuracy for detection responses (‘yes’ vs. ‘no’) was not significantly above chance in any of the brain regions. Similarly, no brain region survived correction for multiple comparisons in the cross-classification of discrimination confidence and decision in detection. The two regions that had the highest accuracy for this classification were the pMFC and BA46. Based on the low classification scores in these regions for decision in detection, these high cross-classification scores are most likely to rely on the fact that ‘yes’ responses tended to be accompanied by high confidence ratings, and not by a true representation of stimulus visibility.

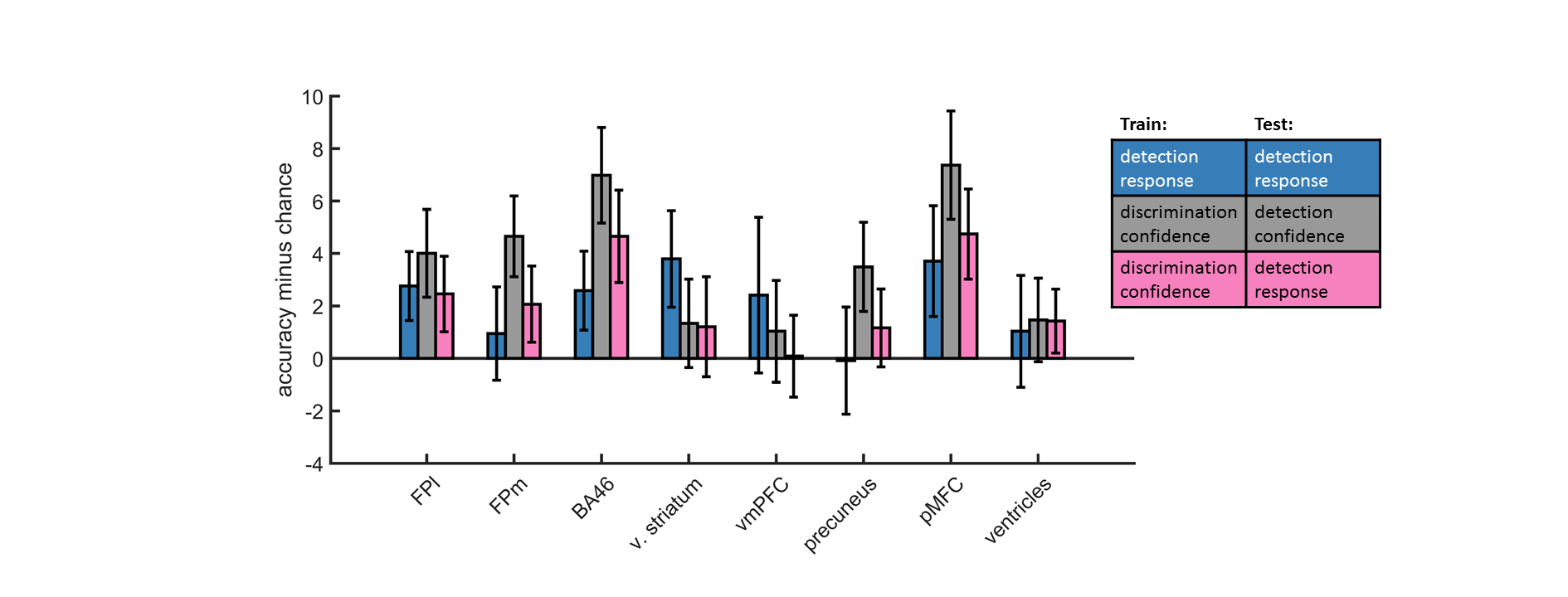


Figure S13: accuracy minus chance for classification of response in detection (‘yes’ vs. ‘no’; blue), and from a cross-classification between tasks: confidence in detection and confidence in discrimination (gray), and confidence in discrimination and decision in detection (pink). Presented cross classification scores are the mean of cross classification accuracies in both directions.