bioRxiv preprint doi: https://doi.org/10.1101/248583; this version posted January 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Finding decodable information that is read out in behaviour

Tijl Grootswagers[1,2,3,*] (tijl.grootswagers@sydney.edu.au)

Radoslaw M. Cichy[4] (rmcichy@zedat.fu-berlin.de)

Thomas A. Carlson[1,2] (thomas.carlson@sydney.edu.au)

[1]School of Psychology, University of Sydney, NSW 2006, Australia

[2]ARC Centre of Excellence in Cognition and its Disorders, 16 University Avenue, NSW 2109, Australia

[3]Department of Cognitive Science, Macquarie University, 16 University Avenue, NSW 2109, Australia

[4]Department of Education and Psychology, Freie Universität Berlin, Habelschwerdter Allee 45, 14195, Berlin, Germany

*corresponding author:

School of Psychology, University of Sydney, NSW 2006, Australia

tijl.grootswagers@sydney.edu.au

Keywords: MVPA, decoding, fMRI, behaviour, visual object recognition

bioRxiv preprint doi: https://doi.org/10.1101/248583; this version posted January 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Abstract

Multivariate decoding methods applied to neuroimaging data have become the standard in cognitive neuroscience for unravelling statistical dependencies between brain activation patterns and experimental conditions. The current challenge is to demonstrate that information decoded as such by the experimenter is in fact used by the brain itself to guide behaviour. Here we demonstrate a promising approach to do so in the context of neural activation during object perception and categorisation behaviour. We first localised decodable information about visual objects in the human brain using a spatially-unbiased multivariate decoding analysis. We then related brain activation patterns to behaviour using a machine-learning based extension of signal detection theory. We show that while there is decodable information about visual category throughout the visual brain, only a subset of those representations predicted categorisation behaviour, located mainly in anterior ventral temporal cortex. Our results have important implications for the interpretation of neuroimaging studies, highlight the importance of relating decoding results to behaviour, and suggest a suitable methodology towards this aim.

bioRxiv preprint doi: https://doi.org/10.1101/248583; this version posted January 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Significance statement

Brain decoding methods are a powerful way to analyse neuroimaging data. An implicit assumption in many decoding studies is that when information can be decoded, then the brain is using this information for behaviour. However, this assumption must be explicitly tested. Here, using visual object categorisation as an example, we separately localised decodable information and information that can be used to predict behaviour. Our findings showed that only from a subset of areas that had decodable category information, we could predict observer categorisation reaction times. Our results highlight the distinction between decodable information and information that is suitably formatted for read-out by the brain in behaviour. Our results have critical implications for the interpretation of decoding studies in general.

bioRxiv preprint doi: https://doi.org/10.1101/248583; this version posted January 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

1 Introduction

Multivariate pattern analysis (MVPA), also called brain decoding, are a powerful tool to establish statistical dependencies between experimental conditions and brain activation patterns measured with functional Magnetic Resonance Imaging (fMRI) (1–7). An implicit assumption often made by the experimenters conducting such analyses is that if information can be decoded, then this information is available to the brain to use in behaviour (8, 9). However, this does not have to be the case, as it could be that the decoded information is different from the signal that is relevant for the brain (8, 10). For example, the decodable information could be epiphenomenal, or the information could be used for a different behaviour than stipulated (cf. 11). Thus, a current challenge in cognitive neuroscience is to relate the decoded information to behaviour. Previous work has approached this issue by, for example, correlating classifier performances to behavioural accuracies (10, 12–17). However, this does not establish the link between brain responses and behaviour at the level of individual experimental conditions. Another approach has been to link behaviour and brain patterns through common similarity structure: if two stimuli evoke similar brain responses, and elicit similar behaviour, brain responses and behaviour are to be linked (18–25). While this approach does link behaviour and brain patterns at the level of single experimental conditions, it is unclear whether the link established such carries over to decision making behaviour such as categorisation.

Recently, a novel methodological approach, called the distance-to-bound approach, has been proposed (26, 27). This method promises to form a tight link between brain activity and behaviour by connecting brain responses at the level of individual experimental conditions directly to perceptual decision-making behaviour. The rationale behind the distance-to-bound approach (16, 26–28) is that in a decision-making task, the brain applies a linear read out process (29) to a neural activation space. Similarly, classifiers in decoding studies fit multi-dimensional hyperplanes to

bioRxiv preprint doi: https://doi.org/10.1101/248583; this version posted January 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

74 separate a neural activation space into separate classes. If the neural activation space used for the
75 decision-making task is the same space that is used by the classifier in a decoding study, then the
76 classifier's decision hyperplane can be interpreted as an observer's decision boundary for the
77 decision-making task. Analogous to the prediction of signal detection theory for the univariate
78 case, in the multivariate space, the closer the input is to the decision boundary, the more
79 ambiguous the evidence for the decision (30). Thus, if for a decision task (e.g., categorisation), the
80 brain uses the same information as the MVPA classifier, then the distance to the classifier
81 hyperplane reflects the ambiguity of the evidence for the decision. Ambiguous evidence in turn
82 predicts longer reaction times (31, 32), and therefore distance to the classifier hyperplane is
83 predicted to negatively correlate with reaction time.
84
85 First studies employing the distance-to-boundary approach have demonstrated its promise in a
86 region of interest based approach using fMRI (27), and a time-varying approach using MEG (33,
87 34). Here, we go beyond those studies by using the distance-to-boundary approach to separately
88 localise information that is decodable, and information that is suitably formatted for read out by
89 the brain. We did not limit our analysis to specific regions of interest, but we used a spatially
90 unbiased approach to create maps of where in the brain information can be used to guide
91 behaviour.
92
93 **2 Results**
94 We examined the relationship between decodable information and information that is suitably
95 formatted for read-out by the brain in the context of decodable information about visual objects
96 and object categorisation behaviour.
97

bioRxiv preprint doi: https://doi.org/10.1101/248583; this version posted January 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

To ensure robustness and generality of our results, we analysed in parallel two independent fMRI datasets (35, 36), with different stimulus sets, and in relation to partly overlapping categorisation behaviours. Stimuli for experiment 1 consisted of animate and inanimate objects (Figure 1A). The animate objects could be further divided into faces, bodies, humans and animals. The stimuli for experiment 2 consisted of 118 visual objects (Figure 1C), which could also be divided into 27 animate and 91 inanimate objects. The inanimate objects were further subdivided into tools, food, and transport objects. Overall, this allowed us to investigate the relationship between decodable information from brain activity and categorisation behaviour for six different distinctions: animate versus inanimate, faces versus bodies, human versus animal, tools versus not tools, food versus not food, and transport versus not food. Note that the negative 'not-X' category was defined as all stimuli that did fall into one of the aforementioned classes.

After collecting human categorisation reaction times (Figure 1E) for each of the six categorical distinctions, we determined the relationship between decodable information and behaviour for each of the six distinctions separately as follows (Fig. 1F). First, we determined where information about objects is present in brain patterns using decoding in a standard fMRI searchlight decoding analysis (6, 37). In detail, at each spatial location (voxel) in an fMRI image, a support vector machine (SVM) was used to classify visual object category based on local brain patterns, resulting in a map of classification accuracies.

We then determined the subset of the locations at which brain patterns were suitably formatted for read-out by the brain using the distance-to-bound approach (26) in a second searchlight analysis. Analogous to the decoding analysis, at each voxel, an SVM was trained to classify visual objects. Diverging at this point from the decoding approach we did not test the classifier, but rather obtained the distance for each exemplar to the hyperplane set by the SVM. We then

bioRxiv preprint doi: https://doi.org/10.1101/248583; this version posted January 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

123 correlated those distances to reaction times acquired in separate categorisation tasks (Fig. 1E). A
124 separate correlation was computed for the two categories in the task. Conducted for every fMRI
125 subject (N=15), this resulted in subject-specific maps of voxels containing object information, and
126 maps of voxels with information that is suitably formatted for readout in a categorisation task.
127 Both sets of maps were subjected to inference statistics at the group level using one-sided sign
128 rank tests and thresholded at $p < 0.05$ (fdr-corrected for multiple comparisons across voxels).
129
130 ***** Figure 1 *****
131
132 **2.1 A subset of locations that have decodable information about animacy also had information**
133 **suitably formatted for animacy categorisation behaviour**
134 Animacy is a pervasive and basic object property according to which any object can be classified as
135 animate or inanimate (38). Previous studies have shown that the division of animate versus
136 inanimate objects is reflected in the large-scale architecture of high-level visual areas such as the
137 ventral temporal cortex (VTC) (38–40), However, it has also been shown that animacy can be
138 decoded not only from VTC, but from the whole ventral visual stream (36, 40, 41). Furthermore,
139 categorical object responses have also been found in the dorsal visual stream (42–44) and in
140 frontal areas (45, 46). This prompts the question of where in the visual system object
141 representations are suitably formatted for read-out by the brain for animacy decisions.
142
143 Corroborating previous studies, we found decodable information about animacy in the entire
144 ventral visual stream from the occipital pole to anterior ventral temporal cortex (Figure 2AB, Table
145 1AD, N = 15, one-sided sign-rank test, $p < 0.05$ fdr-corrected). In addition, we found decodable
146 information in dorsal and prefrontal cortex (Figure 2B) in experiment 2 which had full brain
147 coverage. Localising the brain representations suitable to guide animacy categorisation behaviour

bioRxiv preprint doi: https://doi.org/10.1101/248583; this version posted January 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

(using the distance-to-bound approach) revealed convergent evidence across experiments that only a subset of voxels containing decodable information fulfilled this criterion. In detail, distance-RT-correlations for animate objects were strongest in the high-level regions of the ventral and the dorsal stream (Table 1). For inanimate objects, we found no voxels with significant distance-RT-correlations (cf. 27, 34).

***** Figure 2 *****

2.2 The proportion of region-specific representations suitably formatted for behaviour increases along the ventral stream and decreases along the dorsal stream

We next explicitly determined the degree to which representations in single brain regions within the ventral and dorsal streams are suitably formatted for behaviour. For this we parcellated the cortex (Figure 3A) using a probabilistic topographic map of visual processing areas (47). For each region, we calculated the ratio between the number of significant voxels in the decoding analysis and the total number of voxels, so that a high ratio indicates that a large part of a region contains object representations with categorical information. Similarly, we calculated the ratio between the number of significant voxels in the distance-to-bound analysis and the total number of voxels. Here, a high ratio indicates that a large part of a region contains object representations that are suitably formatted for read out in a categorisation task.

In the ventral stream, we observed a systematic increase in ratio with processing stage, from early visual areas to high-level visual areas, with highest ratios in ventral occipital (VO) and parahippocampal (PHC) cortex (Figure 3 B&C). In contrast, in the dorsal stream we observed a systematic decrease of the correlation ratio with processing stage. In addition, significant animacy decoding information was found in similar proportions in the ventral-temporal areas as in lateral-

bioRxiv preprint doi: https://doi.org/10.1101/248583; this version posted January 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

occipital areas, however, the proportion of voxels with information suitable for categorisation was lower in lateral-occipital areas. This is consistent with the notion that while both these regions contain object representations, the VTC contains location-invariant representations which are essential for object categorisation (10, 24, 48, 49).

In sum, these results show that representations along the ventral stream are shaped for optimal read-out of categorical information (40, 49). In contrast, representations in the dorsal stream might be shaped for the read-out in different tasks (17, 43). These results also suggest that intermediate stages along the ventral and dorsal streams may be similar or partly shared, as suggested by the similar ratios of information suitable for read-out.

***** Figure 3 *****

2.3 Decodable information about subordinate categorisation tasks is also suitably formatted for categorisation behaviour

While animacy categorisation may be based on large-scale representational differences in the visual brain (39, 40, 50, 51), subordinate categorisation tasks (e.g., faces, bodies, tools) may depend more on fine grained patterns in focal brain regions (52–54). Here, we tested whether decodable information about subordinate category membership is also suitably formatted for read out in respective categorisation tasks. We tested two subordinate contrasts for experiment 1: face versus body, and human versus animal using the same general procedure as for animacy. We found that both contrasts were decodable (Table 1B-C). We found a significant correlation between distance to the classifier hyperplane and reaction time for faces in the face versus body task (Figure 4A). Of the subordinate categorisation contrasts in experiment 2 (food, transport or tool versus everything else), transport and tool versus everything else were not significantly

bioRxiv preprint doi: https://doi.org/10.1101/248583; this version posted January 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

198 decodable information nor had they significant correlations (Table 1E-F). Food versus not food
199 resulted in significant decodable information, and significant distance-RT correlations were
200 present for this contrast in the 'not food' category (Figure 4B, Table 1G). Taken together, for some
201 subordinate categorisation contrasts that were decodable, we were successful in localising brain
202 patterns suitably formatted for read-out in behaviour.
203
204 ***** Figure 4 *****
205
206 ***** Table 1 *****
207
208 **3 Discussion**
209 **3.1 Summary**
210 The aim of this study was to examine where in the brain decodable information is suitably
211 formatted for read-out by the brain in behaviour. We found that only a subset of information that
212 is decodable could be related more directly to behaviour using the distance-to-bound approach,
213 which argues for a partial dissociation between decodable information and information that is
214 relevant for behaviour. Besides high-level ventral visual cortex (29, 55, 56), mid-level visual stream
215 (V3 & hV4) and dorsal stream (IPS) regions harboured representations that are formatted suitable
216 to guide behaviour.
217
218 **3.2 Dissociating between decodable information and information that is used in behaviour**
219 Our results speak to a current challenge in neuroimaging, which is to show that information visible
220 to the experimenter is in fact read out by the brain. By showing that only a subset of decodable
221 category information correlates with independent observer categorisation RTs, we demonstrate a
222 dissociation between information that is decodable, and information related to behaviour. To

bioRxiv preprint doi: https://doi.org/10.1101/248583; this version posted January 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

illustrate, consider the question about what regions are used by the brain to perform an object animacy categorisation task (40, 55). When only the result of the animacy decoding searchlight would be taken into account, this might be interpreted as meaning that the brain may use animacy information from anywhere in the ventral stream in animacy categorisation. However, when testing this interpretation directly using the distance-RT-correlation results, it becomes clear that object animacy information is represented for read-out in mid- and high-level visual areas only (9, 10).

It is important to note that not finding a correlation between distance to the classifier hyperplane and RT does not imply this information is necessarily irrelevant or epiphenomenal. For example, it could be that the representations are relevant for a task other than the binary categorisation decision that was used here, or that other read-out models provide a better description (e.g., non-linear models (26)). Therefore, the current approach only allows the positive inference on the level of suitability of decoded information in the context of a limited set of tasks, and a limited set of read-out models. While a correlation with behaviour still does not imply that the information is used by the brain, it shows that the information is formatted in a way that is suitable to be used by the brain for decisions (8–10, 56).

3.3 The contribution of ventral and dorsal visual regions to categorisation behaviour

We found that neural representations suitably formatted for behaviour in categorisation were most prominently located in the anterior regions of the VTC. This corroborates previous studies (27, 57–59), and reinforces the tight link between VTC and visual categorisation behaviour.

However, we found that correlations between distance to boundary and RT were not restricted to anterior regions of the VTC, but were also prominent in V3 and hV4 (Figure 2) as well. This

bioRxiv preprint doi: https://doi.org/10.1101/248583; this version posted January 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

suggests that representations in these areas are suitable for read-out. This is consistent with the view that lower level visual features as encoded in mid-level visual regions could aid faster read-out of category information. V4 is thought of as an intermediate stage of visual processing that aggregates lower level visual features into invariant representations (60). It has been proposed that a direct pathway from V4 to areas that facilitate eye movements (e.g., FEF) accounts for fast saccadic categorisation reaction times (61, 62). Similar pathways to decision making areas would allow the brain to exploit visual feature cues to enable fast responses to ecologically important stimuli (58, 63, 64). An alternative possibility is that read out is not happening directly from V4, but its representational structure is shaped by the low-level feature differences in animacy. This structure is then largely preserved when it is communicated to more anterior areas. This would explain similar correlations between distance to the classifier boundary and reaction time in both V4, and more anterior areas of the ventral stream. Both of these accounts are consistent with recent findings that show differential responses for object categories in mid-level visual areas based on visual feature differences, such as in animacy (21, 41), or object size (65). The extent to which shape information contributes to the read-out process could be further investigated by using the approach from this study with a stimulus set that controls for these visual features (21, 66), or by decoding and predicting reaction times from texture stimuli that have similar features (41).

Finally, we found that correlations with RT were also present in parietal areas. We found that information in early parietal areas also correlates with RT. The classical view is that the ventral and dorsal visual streams are recruited for different function (67), with the ventral stream representing object identities ('what'), and the dorsal stream spatial features ('where'). However, areas in the dorsal stream, such as IPS1 and IPS2 have been found to exhibit similar object-selective responses as areas in the ventral stream (17, 42, 68, 69). In according with this, we found similar correlations

bioRxiv preprint doi: https://doi.org/10.1101/248583; this version posted January 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

between distance and RT in mid-level areas in the ventral and dorsal streams. However, our results also showed that the proportion of correlations between distance and RT decreased further in the dorsal stream, while they increased in the ventral stream (Figure 4). This suggests that representations in the ventral and dorsal streams undergo similar transformations at first, and then diverge for different goals.

3.4 Without a task, neural object representations in the VTC are still formatted for read-out in categorisation decisions

In the current study, the fMRI participants performed an orthogonal task at fixation, and were not actively categorising. Despite this, categorisation reaction times could still be predicted from representations in the visual stream. This highlights that, without a task, information in the visual system is represented in a way that is suitable for read out in behaviour (27, 33). In addition, the orthogonal task in the scanner has the advantage that it avoids motor effects and difficulty confounds in the modelling of the BOLD response (70, 71). Future studies might use the distance-to-bound approach with participants actively performing the same task in the scanner, where we predict that areas involved in the decision making and execution processes would contain information that correlates with reaction time. For example, some areas are believed to preferentially represent task-relevant information, such as areas in the prefrontal cortex (72–74), and in the parietal stream (43, 75, 76). In the absence of an animacy categorisation task, one would predict that animacy information would not be strongly represented in these areas. Yet, our results showed that animacy information can be decoded from prefrontal and parietal areas when participants perform an orthogonal task. Thus, only considering the decoding results would suggest that animacy is represented in these areas. However, our results also showed that the animacy information in these areas was not predictive of reaction time. This again highlights the dissociation between information that can be decoded, and information that is suitable for read

bioRxiv preprint doi: https://doi.org/10.1101/248583; this version posted January 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

298 out in behaviour. A prediction that follows from this is that performing an active object
299 categorisation task in the scanner would change the representations in these task-relevant areas
300 so that they become predictive of reaction time (77).
301
302 **3.5 Asymmetric distance-RT-Correlations in binary categorisation tasks**
303 In both experiments, there was evidence for correlations between distance and reaction time for
304 animate stimuli, but none for the inanimate stimuli, consistent with previous work (27, 33, 34). A
305 possible explanation put forward there was that this discrepancy might be the effect of animacy
306 being a negatively defined category (i.e., “not animate”). Under this hypothesis the animacy
307 categorisation task can be performed by collecting evidence for animate stimuli and responding
308 inanimate only when not enough evidence was accumulated after a certain amount of time. Here,
309 we tested a prediction of this hypothesis using contrasts consisting of two positively defined
310 categories, namely human versus animal, and face versus body. We found that for the face versus
311 body contrast there was a distance-RT-correlation only for faces. This goes against the notion of
312 the negative definition of inanimate as the main reason for a lack of correlation. However, it still is
313 possible that observers still treated these tasks as A or NOT A, rather than A or B, while choosing
314 ‘A’ based on the category that is easiest to detect (34). For example, perceptual evidence for a
315 face would be easier to obtain that evidence for a body-part, as faces share low level visual
316 features (62, 78, 79). Thus, while not explicitly specified as a negative category, it could have been
317 treated as such.
318
319 This suggest that the binary categorisation might be an unnatural way of approaching human
320 categorisation behaviour in the real world. Other operationalisations may be better suited to
321 capture behaviour (11). This could be investigated further by assessing behaviour in tasks that do

bioRxiv preprint doi: https://doi.org/10.1101/248583; this version posted January 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

not require binary categorisation, such as picture naming or visual search, and assessing whether in those tasks the distance-to-bound approach can also be used to predict observer behaviour. Still, it is important to note that the binary task matches the brain decoding task performed by the classifier. The above-chance decoding accuracy in the brain decoding task is commonly interpreted as a similar dichotomy in the brain's representation that the brain can use in a decision. However, when only the information in one of the categories (i.e., animals or faces) can be used to predict decision behaviour, as shown here, then this interpretation needs to be revisited.

3.6 Conclusion

In this study, we combined the distance-to-bound approach (26) with a searchlight decoding analysis to find brain areas with decodable information that is suitable for read-out in behaviour. Our results showed that decodable information is not always equally suitable for read-out by the brain in behaviour. This speaks to the current debate in neuroimaging research about whether the information that we can decode is the same information that is used by the brain in behaviour (8).

4 Materials and Methods

The aim of this study was to correlate distance to a classifier's hyperplane to reaction times (RTs) for categorisation. In this section, we describe the stimuli and fMRI data that were obtained from previous experiments (35, 36). Next, we collected categorisation reaction times for those stimuli on Amazon's Mechanical Turk. Finally, we describe the two-step searchlight procedure used to create decoding and correlation maps of areas involved in visual object categorisation.

4.1 Stimuli

We used the stimuli from two experiments that have been published (35, 36). Stimuli for experiment 1 consisted of 92 visual objects, segmented on a white background (Figure 1A). Stimuli

bioRxiv preprint doi: https://doi.org/10.1101/248583; this version posted January 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

consisted of animate and inanimate objects. The animate objects could be further divided into faces, bodies, humans and animals. Inanimate objects consisted of natural (e.g., plants or fruits) and man-made items (e.g., tools or houses). The stimulus set for experiment 2 consisted of 118 visual objects on natural backgrounds, shown in Figure 1C. A small proportion of the objects (27) were animate. The inanimate objects included subcategories such as tools, or food items. In both experiments, participants were presented with the visual object stimuli while performing an orthogonal task at fixation. Stimuli were displayed at 2.9° (Experiment 1) and 4.0° (Experiment 2) visual angle with 500 ms duration. Images were displayed (overlaid with a grey fixation cross) for 500 ms in random order.

4.2 fMRI recordings

The first experiment (35) had high resolution fMRI coverage of the ventral visual stream (Figure 1B) from 15 participants with a 2 mm isotropic voxel resolution. The second experiment (36) had whole brain from 15 participants with a 3 mm isotropic voxel resolution. In both experiments, at the start of a session, structural images were obtained using a standard T_1-weighted sequence. fMRI data were aligned and coregistered to the T1 structural image, and then normalized to a standard MNI template. General linear models were used to compute t-values for each stimulus (92 and 118, respectively) against baseline.

4.3 Reaction time data

We obtained reaction times for the stimuli in multiple different categorisation contrasts (Figure 1A&B). For experiment 1, these were animate versus inanimate, face versus body, and human versus animal. For experiment 2, we tested animate versus inanimate, tool versus not tool, food versus not food, and transport versus not transport. The RTs were collected using Amazons Mechanical Turk (MTurk). For each of the categorisation contrasts, 50 unique participants

bioRxiv preprint doi: https://doi.org/10.1101/248583; this version posted January 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

performed a categorisation task using the same stimuli as were used in collecting the fMRI data. Participants were instructed to "Categorise the images as fast and accurate as possible using the following keys: (z for X, m for Y)", where X and Y would be replaced with the relevant categories (e.g., animate and inanimate) for the contrast. On each trial, an image was presented for 500ms, followed by a black screen until the participant's response (Figure 1C). The presentation order of the stimuli was randomized and stimuli did not repeat. This resulted in 50 reaction time values per exemplar (one for each participant). Each participant's reaction times were z-scored. Next, we computed the median reaction time (across participants) for each exemplar. his resulted in one reaction time value per exemplar, which were used in the rest of the study.

4.4 Searchlight procedure

For each categorisation contrast and subject, we used a searchlight approach (6, 37) to create maps of decoding accuracy and of correlations between distance to the classifier boundary and categorisation reaction times. In contrast to pre-defined ROI's, which are used to test a-priori hypotheses about the spatial origin of information in the brain, the searchlight results in a spatially unbiased map of decodable information. An overview of the approach is presented in Figure 1D.

To create the decoding accuracy maps, we used a standard searchlight decoding approach (4, 6, 80, 81), as implemented in the CoSMoMVPA decoding toolbox (82). For each voxel in the three-dimensional fMRI voxel space, t-values belonging to a spherical cluster (4 voxel radius) centred around this voxel were arranged into a vector for each stimulus. The stimuli vectors were then split up into five sets for cross-validation. A support vector machine (SVM) classifier was trained on four sets and its prediction accuracy was tested on the left out set. This was repeated five times, using each set for testing once. The mean accuracy across these repetitions was stored at the centre voxel of the sphere. Repeating the process over all voxels resulted in a 3D-map of decoding

bioRxiv preprint doi: https://doi.org/10.1101/248583; this version posted January 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

accuracies. The searchlight procedure was conducted for each subject independently, resulting in one accuracy map for each subject for each task.

Next, we created maps of correlations between distance to the classifier hyperplane and observer categorisation reaction times on the same contrast. Vectors of t-values were obtained for each stimulus in the same manner as the first searchlight procedure. Then, an SVM classifier was trained using all the stimuli, effectively fitting a hyperplane that distinguished between vector patterns according to category. We then computed the distance of each exemplar to this decision hyperplane, resulting in one distance value per exemplar. The distance values were rank-order correlated (Spearman's ρ) to the median MTurk reaction times collected in the behavioural experiment for the same exemplars in the analogous categorisation task. The contribution of each category was assessed individually, by performing the correlations separately for the two sides of the categorisation (i.e., one correlation for animate and one for inanimate exemplars). For each categorisation task this resulted in two correlation maps per subject. The maps of decoding accuracy and correlations were assessed for significance at the group level using sign-rank tests for random-effects inference. The results were thresholded at $p<0.05$, using the false discovery rate (FDR; (83)) to correct for multiple comparisons at the voxel level.

4.5 Relating the results to topographical locations of the visual system

For each of the categorisation contrasts, we identified the locations of the significant voxels with respect to ROIs of the visual system. The significant voxels in the decoding maps and correlation maps were compared to probabilistic topographic maps of visual processing areas (47), which represent for each voxel the visual area with the highest probability. A percentage score for each ROI was then computed, reflecting the percentage of voxels in this ROI that were significant at the group level. We obtained a bootstrapped distribution of percentage scores for each ROI by

bioRxiv preprint doi: https://doi.org/10.1101/248583; this version posted January 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

repeating this procedure 10,000 times, while randomly sampling the subjects with replacement and recomputing the group level statistics. We report the 5^{th}, 50^{th} and 95^{th} percentiles of this distribution. This approach allows quantifying the difference between the number of decoding voxels and correlation voxels per visual ROI.

Acknowledgements

This research was supported by an Australian Research Council Future Fellowship (FT120100816) and an Australian Research Council Discovery project (DP160101300) awarded to T.A.C., and a German Research Foundation grant (CI-241/1-1) awarded to R.M.C. The authors acknowledge the University of Sydney HPC service for providing High Performance Computing resources.

References

1. Carlson TA, Schrater P, He S (2003) Patterns of Activity in the Categorical Representations of Objects. *J Cogn Neurosci* 15(5):704–717.
2. Cox DD, Savoy RL (2003) Functional magnetic resonance imaging (fMRI) "brain reading": detecting and classifying distributed patterns of fMRI activity in human visual cortex. *NeuroImage* 19(2):261–270.
3. Haxby JV, et al. (2001) Distributed and Overlapping Representations of Faces and Objects in Ventral Temporal Cortex. *Science* 293(5539):2425–2430.
4. Haynes J-D (2015) A Primer on Pattern-Based Approaches to fMRI: Principles, Pitfalls, and Perspectives. *Neuron* 87(2):257–270.
5. Kamitani Y, Tong F (2005) Decoding the visual and subjective contents of the human brain. *Nat Neurosci* 8(5):679–685.
6. Kriegeskorte N, Goebel R, Bandettini P (2006) Information-based functional brain mapping. *Proc Natl Acad Sci U S A* 103(10):3863–3868.
7. Tong F, Pratte MS (2012) Decoding Patterns of Human Brain Activity. *Annu Rev Psychol* 63(1):483–509.
8. de-Wit L, Alexander D, Ekroll V, Wagemans J (2016) Is neuroimaging measuring information in the brain? *Psychon Bull Rev* 23(5):1415–1428.

bioRxiv preprint doi: https://doi.org/10.1101/248583; this version posted January 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

9. Ritchie JB, Kaplan DM, Klein C (2017) Decoding the Brain: Neural Representation and the Limits of Multivariate Pattern Analysis in Cognitive Neuroscience. *Br J Philos Sci*. doi:10.1093/bjps/axx023.

10. Williams MA, Dang S, Kanwisher NG (2007) Only some spatial patterns of fMRI response are read out in task performance. *Nat Neurosci* 10(6):685–686.

11. Krakauer JW, Ghazanfar AA, Gomez-Marin A, MacIver MA, Poeppel D (2017) Neuroscience Needs Behavior: Correcting a Reductionist Bias. *Neuron* 93(3):480–490.

12. Naselaris T, Kay KN, Nishimoto S, Gallant JL (2011) Encoding and decoding in fMRI. *NeuroImage* 56(2):400–410.

13. Raizada RDS, Tsao F-M, Liu H-M, Kuhl PK (2010) Quantifying the Adequacy of Neural Representations for a Cross-Language Phonetic Discrimination Task: Prediction of Individual Differences. *Cereb Cortex* 20(1):1–12.

14. van Bergen RS, Ji Ma W, Pratte MS, Jehee JFM (2015) Sensory uncertainty decoded from visual cortex predicts behavior. *Nat Neurosci* 18(12):1728–1730.

15. Walther DB, Caddigan E, Fei-Fei L, Beck DM (2009) Natural Scene Categories Revealed in Distributed Patterns of Activity in the Human Brain. *J Neurosci* 29(34):10573–10581.

16. Philiastides MG, Sajda P (2006) Temporal Characterization of the Neural Correlates of Perceptual Decision Making in the Human Brain. *Cereb Cortex* 16(4):509–518.

17. Freud E, Culham JC, Plaut DC, Behrmann M (2017) The large-scale organization of shape processing in the ventral and dorsal pathways. *eLife* 6:e27576.

18. Bracci S, op de Beeck H (2016) Dissociations and Associations between Shape and Category Representations in the Two Visual Pathways. *J Neurosci* 36(2):432–444.

19. Cohen MA, Dennett DC, Kanwisher N (2016) What is the Bandwidth of Perceptual Experience? *Trends Cogn Sci* 20(5):324–335.

20. Mur M, et al. (2013) Human Object-Similarity Judgments Reflect and Transcend the Primate-IT Object Representation. *Front Psychol* 4. doi:10.3389/fpsyg.2013.00128.

21. Proklova D, Kaiser D, Peelen MV (2016) Disentangling Representations of Object Shape and Object Category in Human Visual Cortex: The Animate–Inanimate Distinction. *J Cogn Neurosci*:1–13.

22. Redcay E, Carlson TA (2015) Rapid neural discrimination of communicative gestures. *Soc Cogn Affect Neurosci* 10(4):545–551.

23. Wardle SG, Kriegeskorte N, Grootswagers T, Khaligh-Razavi S-M, Carlson TA (2016) Perceptual similarity of visual patterns predicts dynamic neural activation patterns measured with MEG. *NeuroImage* 132:59–70.

24. Haushofer J, Livingstone MS, Kanwisher N (2008) Multivariate Patterns in Object-Selective Cortex Dissociate Perceptual and Physical Shape Similarity. *PLOS Biol* 6(7):e187.

bioRxiv preprint doi: https://doi.org/10.1101/248583; this version posted January 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

25. Grootswagers T, Kennedy BL, Most SB, Carlson TA (2017) Neural signatures of dynamic emotion constructs in the human brain. *Neuropsychologia*. doi:10.1016/j.neuropsychologia.2017.10.016.

26. Ritchie JB, Carlson TA (2016) Neural Decoding and "Inner" Psychophysics: A Distance-to-Bound Approach for Linking Mind, Brain, and Behavior. *Front Neurosci*:190.

27. Carlson TA, Ritchie JB, Kriegeskorte N, Durvasula S, Ma J (2014) Reaction Time for Object Categorization Is Predicted by Representational Distance. *J Cogn Neurosci* 26(1):132–142.

28. Kiani R, Cueva CJ, Reppas JB, Newsome WT (2014) Dynamics of Neural Population Responses in Prefrontal Cortex Indicate Changes of Mind on Single Trials. *Curr Biol* 24(13):1542–1547.

29. DiCarlo JJ, Cox DD (2007) Untangling invariant object recognition. *Trends Cogn Sci* 11(8):333–341.

30. Green DM, Swets JA (1966) Signal detection theory and psychophysics. *N Y*.

31. Ashby FG (2000) A Stochastic Version of General Recognition Theory. *J Math Psychol* 44(2):310–329.

32. Ashby FG, Maddox WT (1994) A Response Time Theory of Separability and Integrality in Speeded Classification. *J Math Psychol* 38(4):423–466.

33. Ritchie JB, Tovar DA, Carlson TA (2015) Emerging Object Representations in the Visual System Predict Reaction Times for Categorization. *PLoS Comput Biol* 11(6):e1004316.

34. Grootswagers T, Ritchie JB, Wardle SG, Heathcote A, Carlson TA (2017) Asymmetric Compression of Representational Space for Object Animacy Categorization under Degraded Viewing Conditions. *J Cogn Neurosci* 29(12):1995–2010.

35. Cichy RM, Pantazis D, Oliva A (2014) Resolving human object recognition in space and time. *Nat Neurosci* 17(3):455–462.

36. Cichy RM, Pantazis D, Oliva A (2016) Similarity-Based Fusion of MEG and fMRI Reveals Spatio-Temporal Dynamics in Human Cortex During Visual Object Recognition. *Cereb Cortex N Y N 1991* 26(8):3563–3579.

37. Haynes J-D, et al. (2007) Reading Hidden Intentions in the Human Brain. *Curr Biol* 17(4):323–328.

38. Caramazza A, Shelton JR (1998) Domain-Specific Knowledge Systems in the Brain: The Animate-Inanimate Distinction. *J Cogn Neurosci* 10(1):1–34.

39. Kriegeskorte N, et al. (2008) Matching Categorical Object Representations in Inferior Temporal Cortex of Man and Monkey. *Neuron* 60(6):1126–1141.

40. Grill-Spector K, Weiner KS (2014) The functional architecture of the ventral temporal cortex and its role in categorization. *Nat Rev Neurosci* 15(8):536–548.

41. Long B, Yu C-P, Konkle T (2017) A mid-level organization of the ventral stream. *bioRxiv*:213934.

bioRxiv preprint doi: https://doi.org/10.1101/248583; this version posted January 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

42. Konen CS, Kastner S (2008) Two hierarchically organized neural systems for object information in human visual cortex. *Nat Neurosci* 11(2):224–231.

43. Bracci S, Daniels N, op de Beeck H (2017) Task Context Overrules Object- and Category-Related Representational Content in the Human Parietal Cortex. *Cereb Cortex*:1–12.

44. Freedman DJ, Assad JA (2006) Experience-dependent representation of visual categories in parietal cortex. *Nature* 443(7107):85–88.

45. Freedman DJ, Riesenhuber M, Poggio T, Miller EK (2001) Categorical Representation of Visual Stimuli in the Primate Prefrontal Cortex. *Science* 291(5502):312–316.

46. Freedman DJ, Riesenhuber M, Poggio T, Miller EK (2003) A Comparison of Primate Prefrontal and Inferior Temporal Cortices during Visual Categorization. *J Neurosci* 23(12):5235–5246.

47. Wang L, Mruczek REB, Arcaro MJ, Kastner S (2015) Probabilistic Maps of Visual Topography in Human Cortex. *Cereb Cortex* 25(10):3911–3931.

48. Schwarzlose RF, Swisher JD, Dang S, Kanwisher N (2008) The distribution of category and location information across object-selective regions in human visual cortex. *Proc Natl Acad Sci* 105(11):4447–4452.

49. Cichy RM, et al. (2013) Probing principles of large-scale object representation: Category preference and location encoding. *Hum Brain Mapp* 34(7):1636–1651.

50. Carlson TA, Tovar DA, Alink A, Kriegeskorte N (2013) Representational dynamics of object vision: The first 1000 ms. *J Vis* 13(10):1.

51. Downing PE, Chan AW-Y, Peelen MV, Dodds CM, Kanwisher N (2006) Domain Specificity in Visual Cortex. *Cereb Cortex* 16(10):1453–1461.

52. Downing PE, Jiang Y, Shuman M, Kanwisher N (2001) A Cortical Area Selective for Visual Processing of the Human Body. *Science* 293(5539):2470–2473.

53. Downing PE, Peelen MV (2016) Body selectivity in occipitotemporal cortex: Causal evidence. *Neuropsychologia* 83:138–148.

54. Kanwisher N, McDermott J, Chun MM (1997) The Fusiform Face Area: A Module in Human Extrastriate Cortex Specialized for Face Perception. *J Neurosci* 17(11):4302–4311.

55. DiCarlo JJ, Zoccolan D, Rust NC (2012) How Does the Brain Solve Visual Object Recognition? *Neuron* 73(3):415–434.

56. Klein C (2016) Brain regions as difference-makers. *Philos Psychol* 0(0):1–14.

57. Afraz S-R, Kiani R, Esteky H (2006) Microstimulation of inferotemporal cortex influences face categorization. *Nature* 442(7103):692–695.

58. Hong H, Yamins DLK, Majaj NJ, DiCarlo JJ (2016) Explicit information for category-orthogonal object properties increases along the ventral stream. *Nat Neurosci* 19(4):613–622.

bioRxiv preprint doi: https://doi.org/10.1101/248583; this version posted January 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

59. Hung CP, Kreiman G, Poggio T, DiCarlo JJ (2005) Fast Readout of Object Identity from Macaque Inferior Temporal Cortex. *Science* 310(5749):863–866.

60. Riesenhuber M, Poggio T (1999) Hierarchical models of object recognition in cortex. *Nat Neurosci* 2(11):1019–1025.

61. Crouzet SM, Kirchner H, Thorpe SJ (2010) Fast saccades toward faces: Face detection in just 100 ms. *J Vis* 10(4):16.

62. Honey C, Kirchner H, VanRullen R (2008) Faces in the cloud: Fourier power spectrum biases ultrarapid face detection. *J Vis* 8(12):9–9.

63. Kirchner H, Thorpe SJ (2006) Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited. *Vision Res* 46(11):1762–1776.

64. Thorpe S, Fize D, Marlot C (1996) Speed of processing in the human visual system. *nature* 381(6582):520–522.

65. Long B, Konkle T, Cohen MA, Alvarez GA (2016) Mid-level perceptual features distinguish objects of different real-world sizes. *J Exp Psychol Gen* 145(1):95.

66. Kaiser D, Azzalini DC, Peelen MV (2016) Shape-independent object category responses revealed by MEG and fMRI decoding. *J Neurophysiol* 115(4):2246–2250.

67. Ungerleider LG, Mishkin M (1982) Two cortical visual systems. *Anal Vis Behav*:549–586.

68. Sereno AB, Maunsell JHR (1998) Shape selectivity in primate lateral intraparietal cortex. *Nature* 395(6701):500–503.

69. Silver MA, Kastner S (2009) Topographic maps in human frontal and parietal cortex. *Trends Cogn Sci* 13(11):488–495.

70. Woolgar A, Golland P, Bode S (2014) Coping with confounds in multivoxel pattern analysis: What should we do about reaction time differences? A comment on Todd, Nystrom & Cohen 2013. *NeuroImage* 98:506–512.

71. Hebart MN, Baker CI (2017) Deconstructing multivariate decoding for the study of brain function. *NeuroImage*. doi:10.1016/j.neuroimage.2017.08.005.

72. Woolgar A, Jackson J, Duncan J (2016) Coding of Visual, Auditory, Rule, and Response Information in the Brain: 10 Years of Multivoxel Pattern Analysis. *J Cogn Neurosci* 28(10):1433–1454.

73. Jackson J, Rich AN, Williams MA, Woolgar A (2016) Feature-selective Attention in Frontoparietal Cortex: Multivoxel Codes Adjust to Prioritize Task-relevant Information. *J Cogn Neurosci* 29(2):310–321.

74. Duncan J (2001) An adaptive coding model of neural function in prefrontal cortex. *Nat Rev Neurosci* 2(11):820–829.

75. Freedman DJ, Assad JA (2016) Neuronal Mechanisms of Visual Categorization: An Abstract View on Decision Making. *Annu Rev Neurosci* 39(1):129–147.

bioRxiv preprint doi: https://doi.org/10.1101/248583; this version posted January 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

593 76. Jeong SK, Xu Y (2016) Behaviorally Relevant Abstract Object Identity Representation in the
594 Human Parietal Cortex. *J Neurosci* 36(5):1607–1619.

595 77. Bugatus L, Weiner KS, Grill-Spector K (2017) Task alters category representations in
596 prefrontal but not high-level visual cortex. *NeuroImage* 155:437–449.

597 78. Crouzet SM, Thorpe SJ (2011) Low-Level Cues and Ultra-Fast Face Detection. *Front Psychol* 2.
598 doi:10.3389/fpsyg.2011.00342.

599 79. Wu C-T, Crouzet SM, Thorpe SJ, Fabre-Thorpe M (2015) At 120 msec You Can Spot the Animal
600 but You Don't Yet Know It's a Dog. *J Cogn Neurosci* 27(1):141–149.

601 80. Pereira F, Mitchell T, Botvinick M (2009) Machine learning classifiers and fMRI: A tutorial
602 overview. *NeuroImage* 45(1, Supplement 1):S199–S209.

603 81. Grootswagers T, Wardle SG, Carlson TA (2017) Decoding Dynamic Brain Patterns from Evoked
604 Responses: A Tutorial on Multivariate Pattern Analysis Applied to Time Series Neuroimaging
605 Data. *J Cogn Neurosci* 29(4):677–697.

606 82. Oosterhof NN, Connolly AC, Haxby JV (2016) CoSMoMVPA: Multi-Modal Multivariate Pattern
607 Analysis of Neuroimaging Data in Matlab/GNU Octave. *Front Neuroinformatics*:27.

608 83. Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and Powerful
609 Approach to Multiple Testing. *J R Stat Soc Ser B Methodol* 57(1):289–300.

610

611 **Figure and table legends**

612

613 **Figure 1. General experimental rationale.** Stimuli (**A,C**) used to map fMRI brain responses and
614 brain coverage (**B,C**) for fMRI study 1 and 2 respectively. **E**. Acquisition of reaction times on object
615 categorisation tasks. Reaction times for categorisation contrasts were collected in a different pool
616 of participants than the ones participating in the fMRI experiment. On each trial, a stimulus was
617 displayed for 250ms, and participants categorised it into two categories (exemplarily here:
618 animate vs inanimate) by pressing one of two keys. **F.** The two-partite approach to separately
619 localize decodable information and information that is suitable for read out in behaviour. For both
620 parts, a local cluster of neighbouring voxels (i.e., searchlight) was used to train a linear support
621 vector machine (SVM) on an image category classification task (e.g., animacy). To localize
622 decodable information, the classifier was tested on left-out data, storing the classification

bioRxiv preprint doi: https://doi.org/10.1101/248583; this version posted January 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

accuracy at the centre voxel of the searchlight. To localise information that was suitably formatted for read-out in a categorisation task, the distances of objects to the classifier hyperplane were correlated with the reaction times for the same object images on the same classification task. Repeated for every voxel, this resulted for each subject in one map of decoding accuracies and one of correlations. For visualisation, significant correlation voxels were superimposed on significant decoding accuracy voxels, each showing group average values in significant voxels.

Figure 2. Relationship between decodable information and categorisation behaviour for animacy. A. In experiment 1, decodable animacy information (hot colours) was found throughout the ventral stream. A correlation between distance to the classifier boundary and reaction time for animate stimuli (cool colours) was found in a subset of these areas. **B.** The results of the analysis for experiment 2 corroborated these findings, and showed decodable information in prefrontal areas and in the dorsal visual stream. Correlations between distance and reaction time were also present in the dorsal stream. Visualizations show average results across subjects in significant voxels (N=15, sign-rank test, $p<0.05$ fdr-corrected) projected onto axial slices of a standard T_1 image in MNI space.

Figure 3. Quantifying the decodable information in visual areas and their contribution to categorisation behaviour. A. Locations of topographical ROIs of the visual system (47), containing early visual cortex (EVC) areas V1 and V2, mid-level areas V3 and hV4, high level ventral occipital (VO) and parahippocampal cortex (PHC), temporal occipital (TO) and lateral occipital (LO) areas, areas in the intra-parietal sulcus (IPS), the superior parietal lobule (SPL), and the frontal eye fields (FEF). **B-C.** The ratio between significant voxels in an ROI and the size of the ROI. Orange points show the ratio of voxels within the ROI that had significant animacy decoding performance. Blue points show the ratio of voxels with a significant correlation between distance to the hyperplane

bioRxiv preprint doi: https://doi.org/10.1101/248583; this version posted January 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

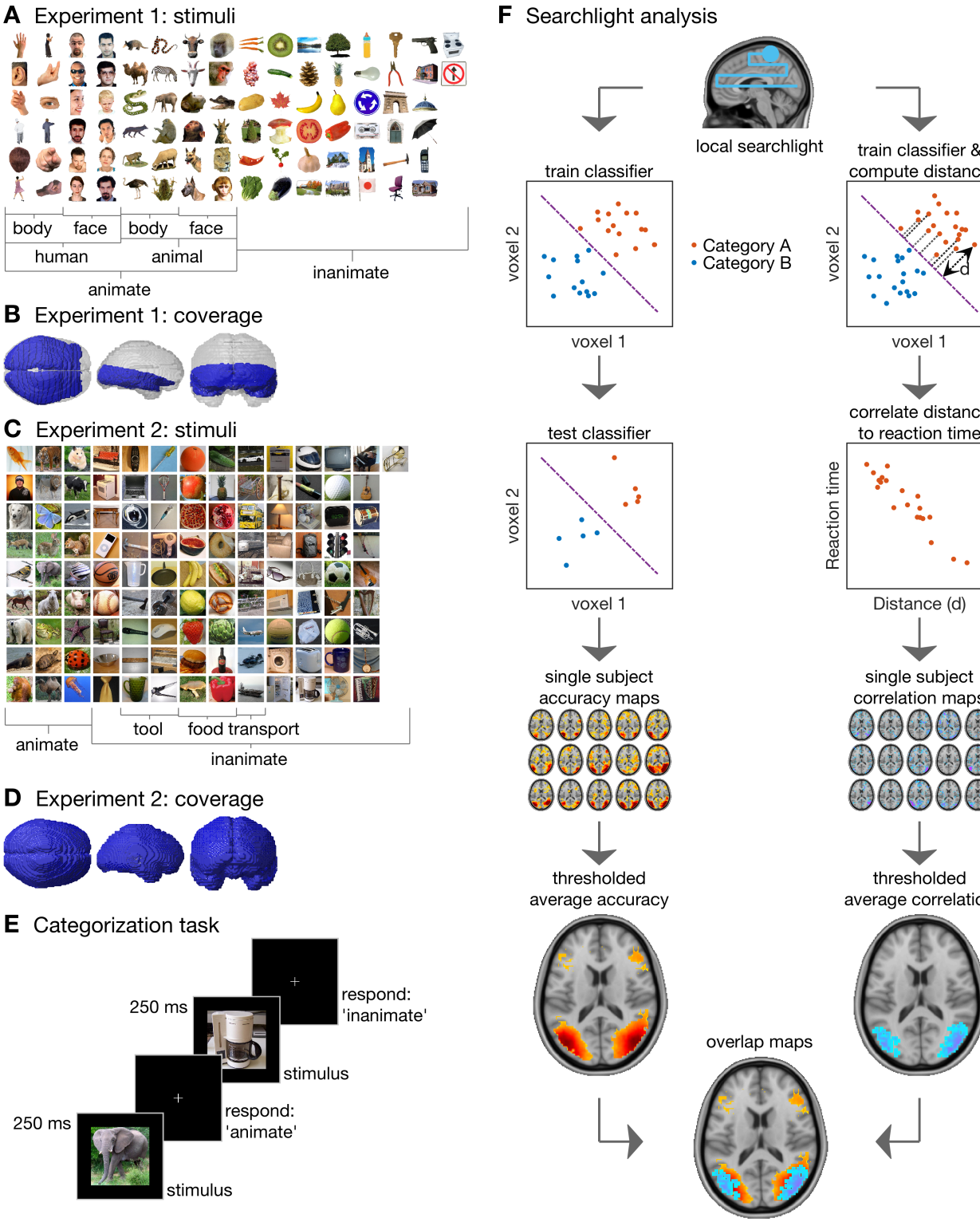
and RT for 'animate'. The lower, middle and upper points on these lines indicate 5^{th}, 50^{th}, and 95^{th} percentiles (bootstrapping of participants 10,000 times). These results quantify the increasing contribution of early to late areas in the ventral visual stream to animacy categorisation behaviour.

Figure 4. Relationship between decodable information and categorisation behaviour for face versus body and food versus not food. A. In experiment 1, decodable face versus body information (hot colours) was found in the entire ventral stream. A distance-RT-correlation for the face stimuli (cool colours) was found in a subset of these areas. **B.** In experiment 2, food versus not food was decodable in some areas in the ventral visual stream. A distance-RT-correlation for the 'not food' stimuli was found in a subset of these areas. Visualisations show average results across subjects in significant voxels (N=15, sign-rank test, $p<0.05$ fdr-corrected) projected onto axial slices of a standard T_1 image in MNI space.

Table 1. Results for all categorisation contrasts. For all categorisation contrasts, we report the number of significant voxels (after correction for multiple comparisons), its peak value (maximum for decoding or minimum for distance-RT-correlation) and uncorrected p-value, and the peak's location in MNI-XYZ coordinates.

bioRxiv preprint doi: https://doi.org/10.1101/248583; this version posted January 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

667 **Figure 1**

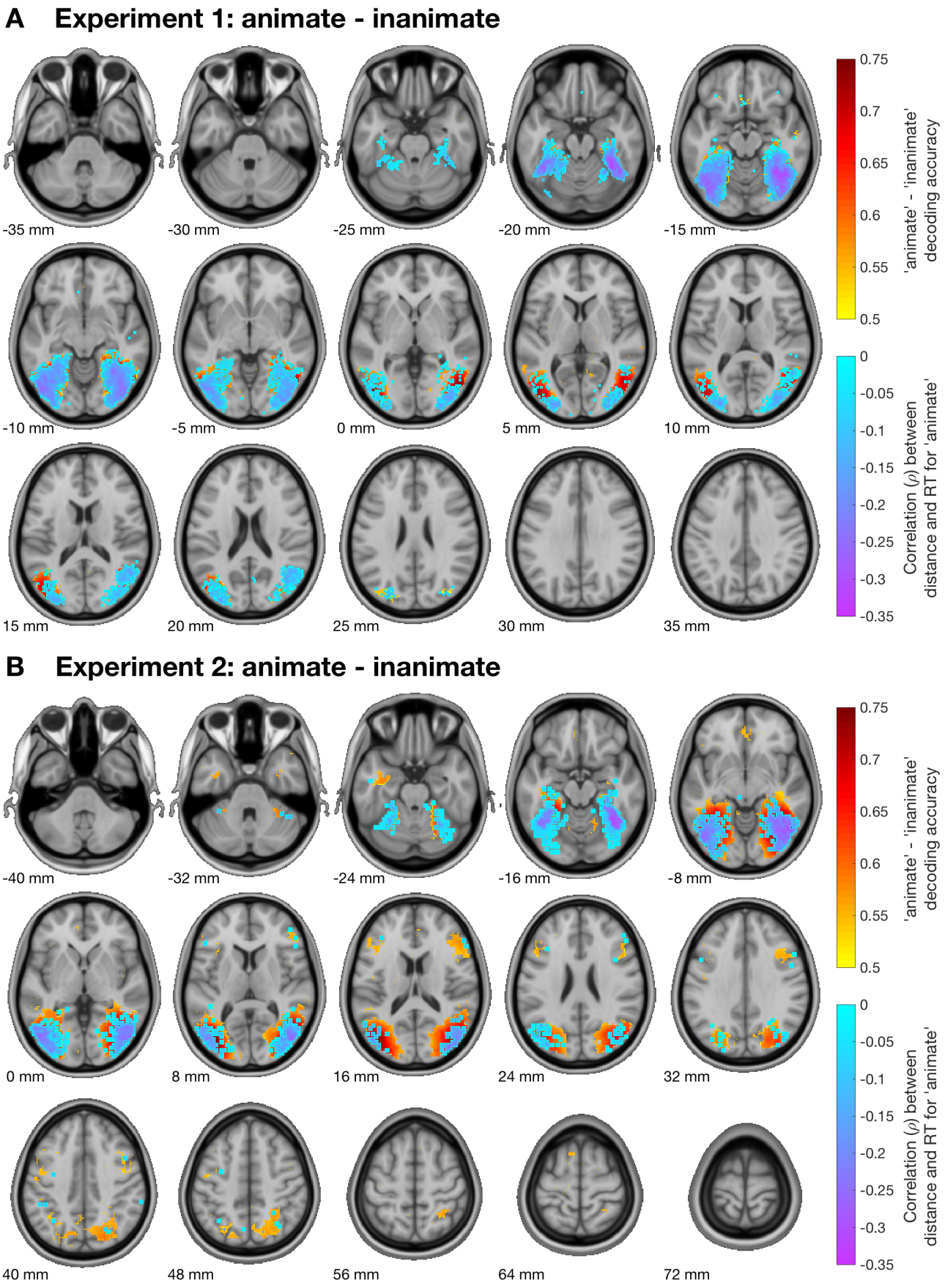


668

669

bioRxiv preprint doi: https://doi.org/10.1101/248583; this version posted January 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

670 **Figure 2**

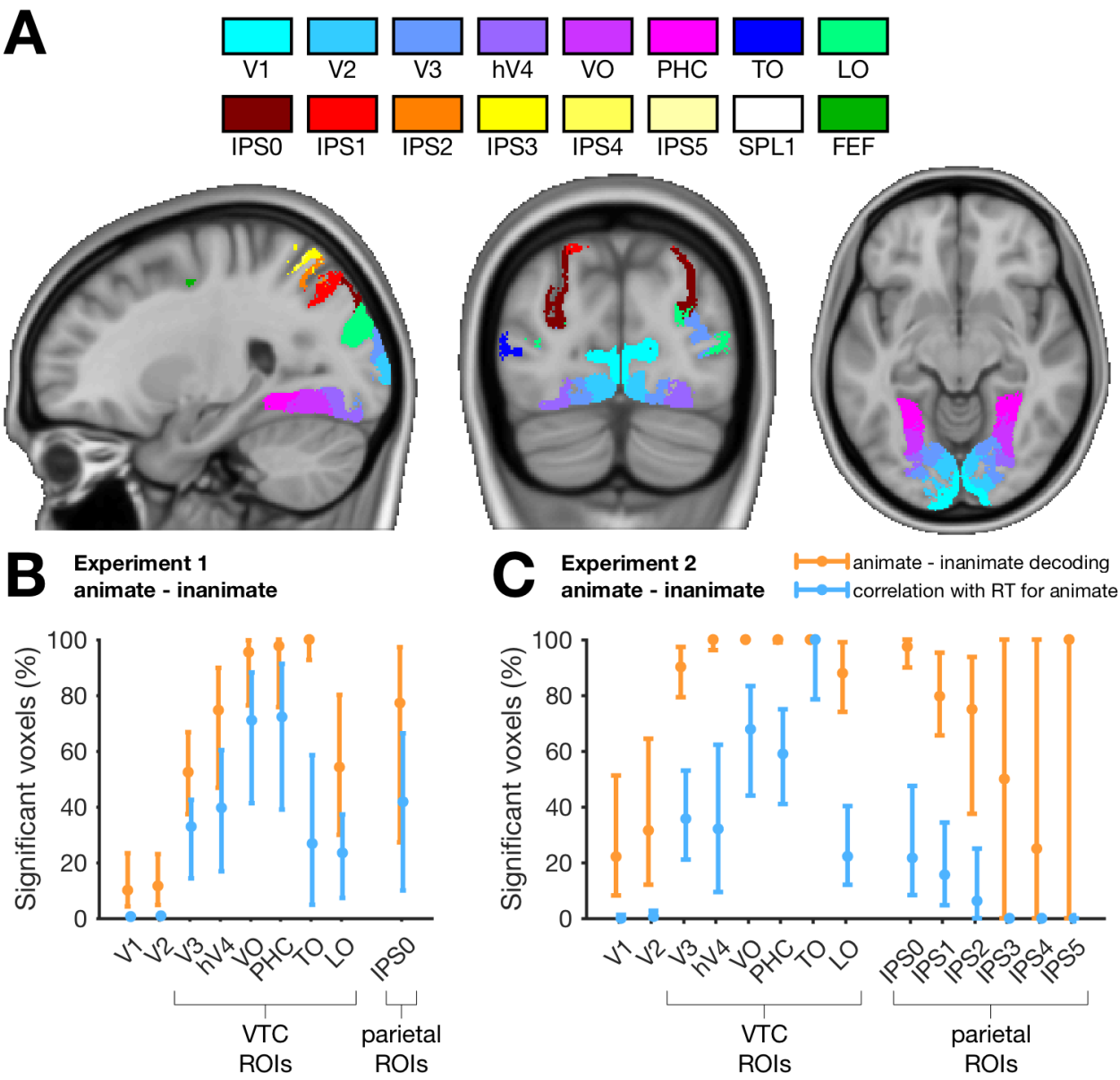


671

672

bioRxiv preprint doi: https://doi.org/10.1101/248583; this version posted January 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

673 **Figure 3**

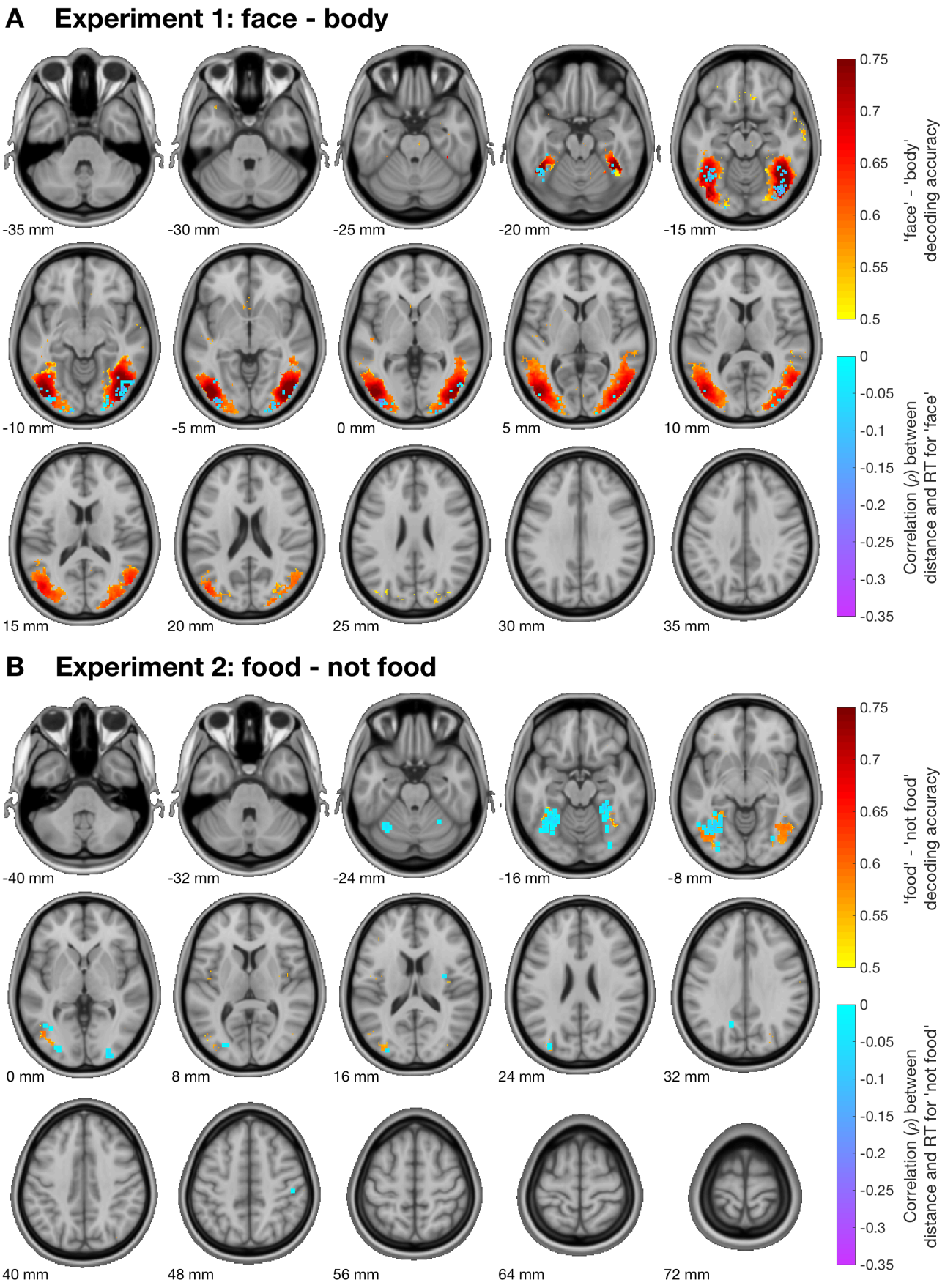


674

675

bioRxiv preprint doi: https://doi.org/10.1101/248583; this version posted January 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

676 **Figure 4**



677

678

bioRxiv preprint doi: https://doi.org/10.1101/248583; this version posted January 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

679 **Table 1**

Contrast	#significant voxels	Max/min	uncorr. p-value	X	Y	Z
A) decoding 'animate' vs 'inanimate' (exp 1)	11745	0.80	0.0000	36	-52	-15
Distance-RT-correlation 'animate'	6410	-0.38	0.0000	38	-58	-19
Distance-RT-correlation 'inanimate'	0	-0.16	0.0021	-48	-58	5
B) decoding 'human' vs 'animal'	4863	0.69	0.0000	22	-90	-13
Distance-RT-correlation 'human'	0	-0.29	0.0002	30	-58	-15
Distance-RT-correlation 'animal'	0	-0.17	0.0072	48	-46	-4
C) decoding 'face' vs 'body'	10661	0.84	0.0000	44	-78	-10
Distance-RT-correlation 'face'	226	-0.32	0.0000	40	-76	-15
Distance-RT-correlation 'body'	0	-0.20	0.0000	-54	-68	16
D) decoding 'animate' vs 'inanimate' (exp 2)	8824	0.80	0.0000	36	-55	-11
Distance-RT-correlation 'animate'	2015	-0.34	0.0000	51	-73	-2
Distance-RT-correlation 'inanimate'	0	-0.12	0.0002	-21	-43	-2
E) decoding 'tool' vs 'not tool'	0	0.58	0.0002	-30	-94	7
Distance-RT-correlation 'tool'	0	-0.25	0.0013	-33	-13	19
Distance-RT-correlation 'not tool'	0	-0.24	0.0001	-33	-52	-17
F) decoding 'transport' vs 'not transport'	0	0.59	0.0001	33	-94	1
Distance-RT-correlation 'transport'	0	-0.32	0.0003	15	50	4
Distance-RT-correlation 'non-transport'	0	-0.18	0.0000	-33	-55	-14
G) decoding 'food' vs 'not food'	1092	0.62	0.0002	36	-55	-14
Distance-RT-correlation 'food'	0	-0.16	0.0177	-18	26	-5
Distance-RT-correlation 'not food'	154	-0.13	0.0002	27	-40	-14

680