

OPEN

Using neural distance to predict reaction time for categorizing the animacy, shape, and abstract properties of objects

J. Brendan Ritchie  & Hans Op de Beeck

A large number of neuroimaging studies have shown that information about object category can be decoded from regions of the ventral visual pathway. One question is how this information might be functionally exploited in the brain. In an attempt to help answer this question, some studies have adopted a neural distance-to-bound approach, and shown that distance to a classifier decision boundary through neural activation space can be used to predict reaction times (RT) on animacy categorization tasks. However, these experiments have not controlled for possible visual confounds, such as shape, in their stimulus design. In the present study we sought to determine whether, when animacy and shape properties are orthogonal, neural distance in low- and high-level visual cortex would predict categorization RTs, and whether a combination of animacy and shape distance might predict RTs when categories crisscrossed the two stimulus dimensions, and so were not linearly separable. In line with previous results, we found that RTs correlated with neural distance, but only for animate stimuli, with similar, though weaker, asymmetric effects for the shape and crisscrossing tasks. Taken together, these results suggest there is potential to expand the neural distance-to-bound approach to other divisions beyond animacy and object category.

Using multivariate pattern analysis (MVPA) several neuroimaging studies have investigated object category-selectivity in regions of the ventral visual pathway¹. At the same time, some findings have been interpreted as showing that apparent categorical effects in high-level regions of the ventral pathway may be accounted for by more low-level visual features of stimuli^{2–4}. Only recently have studies begun to employ stimulus designs that tease apart the separate contribution of object category and other visual properties, such as shape, to the neural response in ventral pathway brain regions^{5–7}. A virtue of the stimulus designs of these studies is that, by making category and shape orthogonal properties of object images, they help rule out the possibility that the categorical effects in later regions of the ventral pathway are accounted for by low-level visual features of stimuli⁸.

Besides possible visual confounds, another issue is that evidence of category-related information in a brain region does not guarantee that it is functionally exploited and therefore represented. A first step in trying to address this issue is to show that decodable information can be used to predict behavior^{9–11}. One way of moving beyond this first step is to attempt to not just predict, but also partially explain, observer behavior using neural distance-to-bound analysis (NDBA), which involves correlating distance from a classifier decision boundary in activation space with observer categorization reaction times (RTs). Based on the predictions of classic psychophysics^{12,13}, if the neural activation space reflects the representation that determines observer categorization response, then RTs should negatively correlate with distance from the decision boundary¹⁴. Recent studies have used NDBA to predict observer RTs for animacy categorization tasks based on neural distances measured with human fMRI and MEG^{15–18}.

To date, none of the studies employing NDBA have used stimulus designs that dissociate object category from other visual features of object images such as shape, or used NDBA to determine whether neural distance predicts observer performance when discriminating real-world shape. In the present study we therefore sought to determine whether, when animacy and shape properties of objects are orthogonal, neural distance would still predict observer RTs on both the animacy and shape categorization tasks, and whether the relationship between

Laboratory of Biological Psychology, Department of Brain and Cognition, KU Leuven, 3000, Leuven, Belgium. Correspondence and requests for materials should be addressed to J.B.R. (email: brendan.ritchie@kuleuven.be)

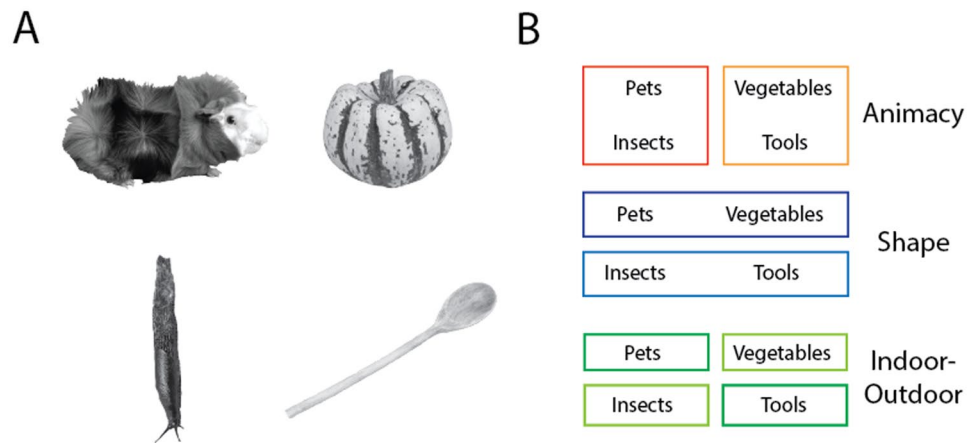


Figure 1. Stimuli and categorization task design. **(A)** Representative natural images of the pets, insects, vegetables, and tools stimuli. The sources for 3/4 of the displayed images are as follows: guinea pig (<https://commons.wikimedia.org/wiki/File:AniarasKelpoKalle.jpg>); slug (<https://www.flickr.com/photos/brewbooks/2606728819/>); and wooden spoon (https://upload.wikimedia.org/wikipedia/commons/7/7b/Wooden_Spoon.jpg). These 3/4 images are published in compliance with a CC BY-SA license (<https://creativecommons.org/licenses/by-sa/3.0/>). The full set of stimulus images used in the study can be found at: <https://osf.io/3dpqe/>. **(B)** the category groupings of the stimuli for the animacy, shape, and indoor-outdoor tasks.

neural distance and task might vary between low- and h-level visual cortex. In previous applications of NDBA, it has typically been assumed that representations of object categories reflect distinct, and linearly separable, areas of the activation space in a brain region. This raises the question of whether neural distance might predict RTs when categories are not obviously linearly separable in activation space. In our stimulus design object categories were linearly separable along the dimensions of animacy and shape. Thus, we also sought to determine whether a combination of animacy and shape distances might predict RTs when task categories crisscrossed the two stimulus dimensions, and so were not linearly separable along these dimensions.

Materials and Methods

Participants. 15 adult volunteers (8 Female; mean age = 30; right-handed) participated in the experiment. All participants had normal or corrected vision, provided written informed consent, and were financially compensated for their participation. The experiment was approved by the Medical Ethics Committee of UZ/KU Leuven and all methods were performed in accordance with the relevant guidelines and regulations.

Stimuli. Stimuli were 32 natural images of objects, converted to greyscale with scene background removed (Fig. 1A). All images were cropped to 700×700 pixels and subtended $\sim 10^\circ$ of visual angle in the scanner. Images were selected to include two clusters each of animate and inanimate objects, which were counterbalanced into two clusters of real-world shape: those with a high-aspect ratio, or “bar-like” shape, and those with a low-aspect ratio, or “blob-like”, shape (Fig. 1B). This resulted in 4 subordinate clusters of 8 images each: pets, insects, tools, and vegetables. Stimulus presentation and control for all experiments in the study was via PC computers running the Psychophysical Toolbox package¹⁹, along with custom code, in Matlab (The Mathworks).

Image analysis. To assess whether low-level image properties predicted stimulus animacy or shape, we applied the GIST model, which provides a summary of local orientation and spatial frequency information in an image, to our stimuli²⁰. Each image was segmented into a 4×4 grid, and Gabor filters with 8 different orientations and 4 spatial frequencies were applied independently to each grid block of the image. This procedure resulted in a 512 value descriptor vector ($16 \text{ image blocks} \times 8 \text{ orientations} \times 4 \text{ spatial frequencies}$) for each image. We then applied a linear discriminant analysis (LDA) classifier to the descriptor vectors to carry out cross-decoding for the shape and animacy labels of the images. For example, a classifier trained to discriminate the animacy of (shape-matched) insects and tools images was tested on the (shape-matched) pets and vegetable images, and vice versa. This analysis allowed us to assess whether any information about animacy category or real-world shape generalized across stimulus clustering along the orthogonal dimension.

Categorization tasks. Subjects performed three tasks in the scanner: an animacy task (is the object in the image “capable of self-movement”); a shape task (is it more “bar-like or blob-like”); and a contrived “indoor-outdoor” task, which crisscrossed the animacy and shape dimensions (Fig. 1B). For the indoor-outdoor task it was emphasized to participants that one feature of some of the objects is that they tend to be found or kept indoors. Specifically, the tool stimuli were kitchen and bathroom items, and the pets (e.g. guinea pigs) tend to be kept in cages indoors. In contrast, we highlighted that the vegetables in our images are typically grown outdoors in the garden, and that the insects in our images are also found outside in trees or gardens. Thus, for the indoor-outdoor task, subjects were instructed to use a binary rule for categorizing the object images into groups that crisscrossed the animacy and shape divisions, but without overt reference to a nonlinear decision rule. All

three tasks were explained to subjects prior to scanning. Dependent measures for these tasks were observer choice responses and reaction times.

Scanning procedure. Scanning consisted of a single sessions of 9 functional scans, followed by an anatomical scan. Using a rapid event-related design, each run consisted of a random sequence of trials including two repeats of each of the 32 stimuli and 11 fixation trials, for a total of 86 trials per run. Each stimulus trial began with the stimulus being presented for 1000 ms, followed by 2000 ms of fixation. The duration of each run was 4m 40 s. During scanning, subjects first performed either the animacy then shape, or shape then animacy task, followed by the indoor-outdoor task, for 3 runs per task. Responses were made using the middle and index finger of the right hand. Since stimulus durations were longer than is typical when collecting RT data, participants were instructed to respond while the stimulus was still on the screen. To control for possible motor confounds, the response-button mapping was counterbalanced between runs.

Imaging parameters. Data acquisition was carried out using a 3T Philips scanner, with 32-channel coil, at the Department of Radiology of the KU Leuven university hospitals. Functional MRI volumes were acquired using a 2D T2*-weighted echo planar (EPI) sequence: TR = 2 s; TE = 30 ms; FA = 90 deg; FoV = 216; voxel size = 3 × 3 × 3 mm; matrix size = 72 × 72. Each volume consisted of 37 axial slices aligned to encompass as much of the cortex as possible, with no gap. The T1-weighter anatomical volumes were acquired with an MP-RAGE sequence, 1 × 1 × 1 mm resolution.

fMRI preprocessing and analysis. Preprocessing and analysis of the MRI data was carried out using SPM12 (v. 6906). For each participant, fMRI data was slice-time corrected, motion corrected, coregistered to the individual anatomical scan, normalized to standard MNI space, and then smoothed using a Gaussian kernel, 6 mm FWHM²¹. After preprocessing the BOLD signal for each stimulus, at each voxel, was modeled using a GLM. The predictors for the model consisted of the 32 stimulus conditions and 6 motion correction parameters (translation and rotation along the *x*, *y*, and *z* axes). The time course for each predictor was characterized as a boxcar function (1000 ms duration) convolved with the canonical hemodynamic response function, with onsets consisting of the onset times for the two repeats of a stimulus condition from each run. The GLM analysis produced one parameter estimate for each voxel, for each stimulus predictor, for each run.

Decoding searchlight analysis. To determine where information about category and shape might be located, we carried out a volume-based decoding searchlight analysis using LDA classifiers, as implemented in the CoSMoMvPA toolbox²². In this analysis, the total brain volume is partitioned into distinct “neighborhoods”, each one centered on a single brain voxel and defined to include 100 adjacent voxels based on their 3D image coordinates. For each neighborhood, an LDA classifier was trained to discriminate pattern responses across the 100 voxels to determine whether they carried information that discriminated between task conditions, such as whether a voxel pattern was produced by an animate or inanimate exemplar image²³. The classifiers were trained using two different cross-validation procedures. First, we utilized standard leave-one-run-out cross-validation. Second, mirroring the image analysis, we performed a cross-decoding searchlight analysis⁶. To control for multiple comparisons, all *p*-values were FDR-adjusted. To isolate neural features for carrying out NDBA, we selected the 10% highest classifying voxels within the area that showed significant decoding (see below).

Neural distance-to-bound analysis. NDBA can be motivated by considering how LDA classifiers assign category labels to test data based on a linear model constructed for the training data. This model projects training data points in the feature space onto a discriminant axis, which maximizes the between-class (while minimizing the within-class) variance, with a decision boundary being positioned orthogonal to the discriminant axis. Training data is then likewise projected onto the discriminant axis and labeled based on its position relative to the decision boundary. In typical applications, the only dependent measure of interest is the classifier accuracy, when averaged across cross-validation folds. However, when used to classifier pattern responses in neural activation space, the relative distance of activity patterns along the discriminant axis from the decision boundary may also contain information about how visual categories are represented in a brain region. NDBA exploits this potential source of information in order to predict observer categorization behavior.

The rationale for NDBA is as follows: according to distance-to-bound models of discrimination and categorization, such as signal detection theory and its multidimensional extension^{12,13}, evidence quality increases with distance from a decision boundary through psychological space (where dimensions may reflect mentally represented properties of stimulus such as size and orientation). Furthermore, RTs tend to decrease as the quality of evidence for a decision increases. Therefore, if differences in the latencies of RTs across stimuli are primarily the result of differences in the quality of the evidence they afford, then RTs for the stimuli should negatively correlate with the distance of stimulus representations from the decision boundary in the psychological space. NDBA applies the same reasoning to neural activation spaces in order to predict behavior¹⁴. If an activation space carries information that is relied on by an observer performing a categorization task, then one can use a linear classifier applied to the space to estimate a decision boundary, and then correlate the distances of the neural responses of stimuli from this boundary with observer RTs for the same stimulus conditions. Thus, following on previous work using NDBA, we correlated observer RTs with neural distances from a classifier decision boundary through activation space.

Besides RTs, we have also found that neural distances negatively correlate with observer accuracies on animacy tasks, and that the drift rate parameter of an evidence accumulation model, which jointly models both dependent variables, can produce a larger correlation with neural distance¹⁸. A simpler method for combining the strengths of both measures into a single value is via an efficiency score, which assumes a linear relationship between choice speed and accuracy. We opted to use the linearly integrated speed-accuracy score (LISAS), which respects this

linear relationship and constitutes a form of accuracy (i.e. proportion error) adjusted RT²⁴. After computing this value for each stimulus for each subject, we group averaged the scores to generate a single behavioral vector to correlate with the individual subject neural distances. As a proxy for neural distances we used decision weights of an LDA classifier, averaged across cross-validation folds with a 4/5 train-test split. These values were calculated from regions identified by the 10% highest classifying voxels as determined by the searchlight analysis for animacy and shape, since RT-distance effect size largely tracks decodability across spatial and temporal features^{16,17}.

Since it has previously been observed that the RT-distance relationship for animacy tasks is primarily driven by animate stimuli we separately correlated the distances for each of the two categories for each of the three tasks¹⁷. Crucially, for the indoor-outdoor task, we separately scaled the neural distances from the animacy and shape classifier boundaries for the high classifying voxels to percentiles, so that both sets of distances would receive equal weight. These distances were then summed, following the rationale of city block distance metrics assuming shape and animacy are independently represented dimensions. These scaled and summed values were then correlated with the adjusted RTs. Thus the neural distances for the indoor-outdoor task reflected information about both animacy and shape, which were crisscrossed dimensions in our stimulus design.

While in previous work we have used rank-order correlations in the present study we linearly correlated the neural distance and adjusted RT vectors, since this allowed us to adopt a joint measure of reliability from psychometrics that has been used in previous electrophysiological and neuroimaging studies in monkeys^{25,26}. For this we computed the mean split-half correlations for both adjusted RTs and neural distances, and then applied the Spearman-Brown formula to estimate the reliability of the full data sets. The joint reliability is then the square root of the product of the two coefficients²⁷. Since these values are positive, but the correlations between adjusted RTs and neural distances are predicted to be negative, the sign-value of the reliability estimates were inverted for visualization.

Results

Behavioral performance on categorization tasks. Participants performed three categorization tasks for the natural image stimuli while lying in an fMRI scanner. For the animacy task participants were highly accurate in their choice responses (mean accuracy = 0.91 ± 0.09), and were less accurate for the shape task (mean accuracy = 0.85 ± 0.11), and especially the indoor-outdoor task (mean accuracy = 0.81 ± 0.14). For RTs participants were fastest at the shape task (median RTs = 818 ms, ± 8 ms), than the animacy task (median RTs = 840 ms, ± 10 ms), and slowest for the indoor-outdoor task (median RTs = 896 ms, ± 10 ms). This pattern of behavioral results is consistent with the fact that subjects reported the indoor-outdoor task to be the most difficult, which was to be anticipated based on the contrived nature of the category division applied to the stimulus images. Notably, the data of one participant, who was at chance for the indoor-outdoor task, was excluded from all further analysis.

In preparation for NDBA, accuracy and RT data were combined. Adjusting RTs using observer accuracy assumes a negative correlation between the two measures²⁴. This was generally observed across tasks, albeit only significantly for the animacy (12/15 participants, mean $r = -0.22 \pm 0.25$, $t(14) = 03.33$, $p < 0.01$) and shape tasks (10/15 participants, mean $r = -0.18 \pm 0.29$, $t(14) = -2.5$, $p < 0.05$), but not the indoor-outdoor task (10/15 participants, mean $r = -0.11 \pm 0.24$, $t(14) = -1.8$, $p = 0.09$).

Decoding searchlight for object animacy and shape. Decoding searchlight analysis was performed to isolate high classifying voxels for object animacy and shape (Fig. 2A). This revealed animacy-related information in the lateral occipital cortex (LOC), and ventral temporal cortex (VTC), but crucially not early visual cortex (EVC). For shape information decodable information was found across the occipital lobe including EVC, as well as LOC and VTC. However, the highest 10% classifying voxels for these analyses did not overlap, and were primarily in LOC for animacy, and EVC for shape. Image analysis showed that a cross-decoding classifier could readily distinguish between shape-types based on GIST descriptors (mean accuracy: 0.88; individual cross-validation fold accuracy: 0.82 and 0.94), but crucially, not animacy (mean accuracy = 0.38; individual cross-validation fold accuracy: 0.38 and 0.38), suggesting low-level properties of the images contained little category-identifying information. In contrast, cross-decoding searchlight analysis revealed significant decoding for both animacy and shape, with significant clusters that tended to overlap with the regions of peak decoding using standard cross-validation (Fig. 2B).

Neural distance predicts categorization behavior across tasks. First, we correlated the adjusted RTs for the animacy task with neural distances from an animacy classifier boundary in LOC and EVC. In LOC there was a robust RT-distance correlation for the animate images (mean $r = -0.58$, $t(14) = -13.02$, $p < 0.001$), but only a borderline significant effect for the inanimate images (mean $r = -0.12$, $t(14) = -2.08$, $p = 0.06$). The difference between the mean correlations for the two categories was also highly significant ($t(14) = -5.40$, $p < 0.001$). Furthermore, the joint measures of reliability (grey bars in Fig. 3A) suggest that there is explainable variance for the inanimate stimuli, but it is not well-captured by neural distance. In EVC there was a small but significant RT-distance correlation (mean $r = -0.13$, $t(14) = -2.58$, $p < 0.05$) for the animate stimuli but not the inanimate stimuli (mean $r = 0.08$, $t(14) = 1.10$, $p = 0.29$), even though animacy could not be decoded from this region.

Second, we also calculated the correlations between adjusted RT neural distance for our shape task (Fig. 3B). There were significant negative correlations for both the blob-like (mean $r = -0.32$, $t(14) = -4.50$, $p < 0.01$) and bar-like stimuli (mean $r = -0.17$, $t(14) = -3.03$, $p < 0.01$) in EVC, where there was greatest shape decoding, with no significant difference between the effects. A significant negative correlation was also observed in LOC for the bar-like stimuli (mean $r = -0.16$, $t(14) = -3.34$, $p < 0.01$), but not for the blob-like category (mean $r = -0.10$, $t(14) = -1.46$, $p = 0.17$), even though the joint reliability for the blob-like images was similar for both ROIs. In

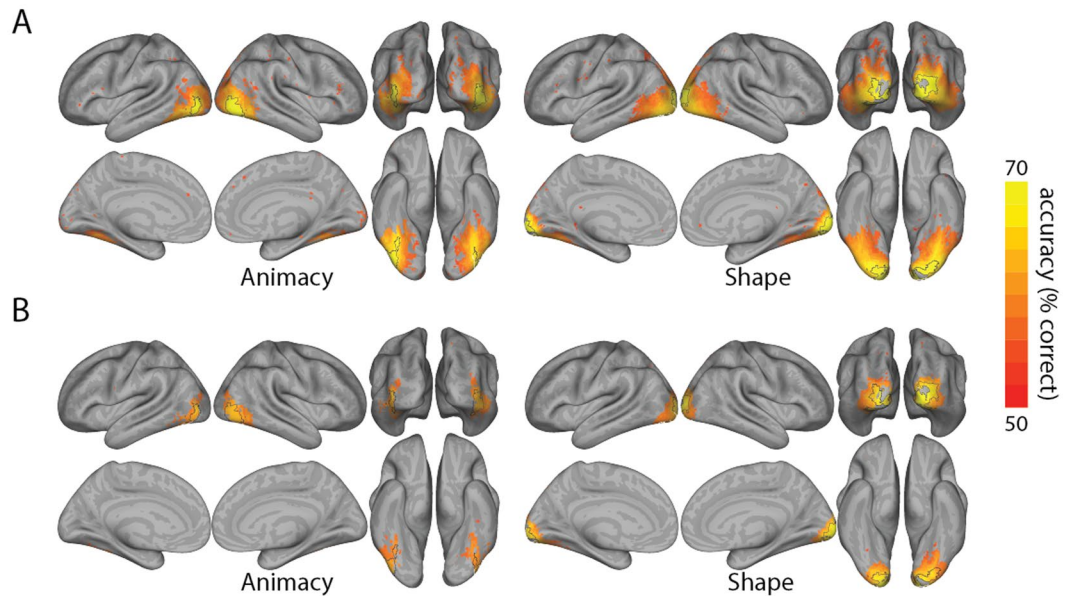


Figure 2. Decoding searchlight results. Results of searchlight analysis using standard leave-one-run-out cross-validation decoding (A), and cross-decoding (B), for both animacy and shape. Accuracy maps are projected onto the inflated cortical surface using CARET³⁷, and thresholded at $p < 0.005$ (FDR-adjusted). Black lines on the surfaces indicate the location of the 10% of voxels with highest classifier accuracy when using standard cross-validation (primarily in LOC for animacy, and in EVC for shape).

contrast, for the bar-like stimuli the correlation in LOC was close to the noise ceiling, which was much lower than in EVC.

Third, we carried out the same analysis for the crisscrossing indoor-outdoor task (Fig. 3C). Here we observed no effects in EVC for either the indoor (mean $r = 0.10$, $t(14) = 1.84$, $p = 0.09$) or outdoor (mean $r = -0.02$, $t(14) = -0.21$, $p = 0.84$) category stimuli. There was an effect for the indoor stimuli in LOC (mean $r = -0.12$, $t(14) = -3.04$, $p < 0.01$) but not the outdoor stimuli (mean $r = 0.00$, $t(14) = -0.01$, $p = 0.99$). With the exception of the one significant effect, the individual data points were highly variable with the mean correlation coefficients even trending in the wrong, positive direction. Overall, the joint reliability was also lower for the indoor-outdoor task than for the other two tasks.

Discussion

In the present study we sought to determine whether neural distance would predict observer RTs when stimuli are balanced along orthogonal object category and shape dimensions. To this end, we correlated neural distances from classifier decision boundaries for object animacy and shape in EVC and LOC with accuracy adjusted RTs from three categorization tasks: animacy, shape, and indoor-outdoor. In each case we observed, for at least one side of the category divisions, the predicted negative correlation between neural distance and adjusted RTs.

Stimulus animacy was highly decodable from LOC, consistent with previous findings^{28–30}. Also, as cross-decoding was used to both investigate whether category information can be discriminated by image properties and to isolate category-specific information in the ventral pathway^{5,6,31}, our results cannot be explained by the sorts of image properties captured by GIST^{2–4}. When neural distances from this region were correlated with adjusted RTs we found a strong RT-distance effect for the animate, but not inanimate, exemplars. This same pattern of results has been observed in previous studies using NDBA to investigate the animacy division. In particular, Carlson *et al.*¹⁵ found a robust RT-distance effect in human VTC, which was driven entirely by the animate exemplars, and taking a searchlight approach to NDBA Grootswagers *et al.*¹⁷ found that neural distance for animate exemplars correlated with RTs in LOC and VTC, and largely tracked the areas of peak decoding. The same asymmetry has also been observed when applying NDBA to human MEG time-series data^{16,18}. The fact that a weak RT-distance effect was also observed for animate images in EVC is also consistent with previous findings of Carlson *et al.*¹⁵, and suggests that the visual properties represented in the region influence how easily a stimulus can be categorized even in situations in which these low-level properties by themselves are not sufficient for categorization.

Methodologically, the present study differed from previous studies using NDBA in three important respects. First, unlike in these previous studies, we controlled shape properties so that overall the animacy RT-distance effects cannot be easily explained by low-level properties⁸. For this reason, our results provide more compelling evidence that observers are recruiting category-related information encoded in high-level visual cortex when performing the animacy categorization task.

Second, previously we have found that observer choice accuracies also correlate with neural distance. When both accuracy and RT distributions are jointly modeled using evidence accumulation models, this provides a more complete description of the observer performance, and the drift rate parameters of these models can also be

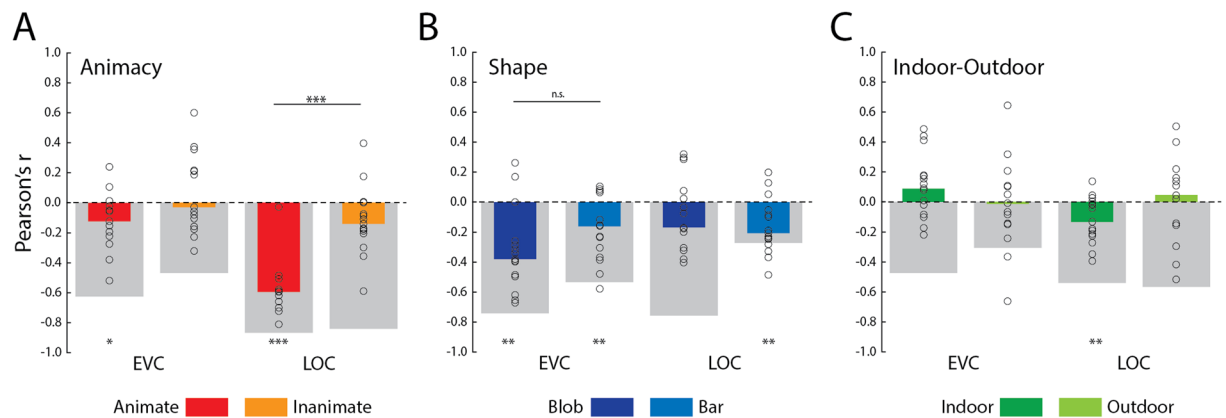


Figure 3. Neural distance-to-bound results. Results of neural distance-to-bound analysis for each of the three categorization tasks: animacy (A); shape (B); and indoor-outdoor (C). Color coding of median bars reflects the stimulus groupings for each task from Fig. 1B. Neural distances are for EVC and LOC features, based on the highest classifying voxels from the searchlight analysis (Fig. 2A). Thick gray bars are the inverted joint reliability of the neural distances and accuracy adjusted RTs. * p , 0.05; ** p < 0.01; *** p < 0.001.

correlated with neural distance¹⁸. Efficiency scores provide a simpler method for combining both measures, and in the present study we used the recently proposed LISAS²⁴. Future studies might systematically compare whether drift rates of evidence accumulation models, which require more intensive procedures for parameter estimation, in fact show superior effects to the analytically simple efficiency scores when carrying out NDBA.

Third, an important consideration in analyses that aim to characterize the structure of activation spaces of brain regions is to have an estimate of data reliability. For example, this is common practice when carrying out representational similarity analysis (RSA)^{32,33}. In particular, cross-validated distances may provide the best technique for estimating the neural dissimilarity relationships at the heart of RSA, and data reliability can be used to estimate a noise ceiling for correlations between model and neural dissimilarity matrices. To date, previous studies using NDBA have not used the same quality controls in estimating neural distances from a decision boundary. In the present study we rectified this by cross-validating our estimates of neural distance and using a joint measure of both behavioral and neural reliability to provide an upper bound on effect sizes^{25,26}. Notably, the same joint measure of reliability could be employed in RSA studies where linear correlation coefficients can be used to compare dissimilarity matrices of neural and behavioral data. The importance of including such reliability estimates in the current study is illustrated by the large variation in reliability even when analyses target the same brain area, as can be seen in Fig. 3. This variation shows, for example, that the relatively low correlation between RTs and neural distance for elongated shapes can mostly be predicted from a low reliability of the data that are correlated. In contrast, the low correlation for inanimate objects is much lower than what would be predicted based upon the reliability of the data.

It is also useful to compare our results to recent studies using stimulus designs that orthogonalized object category and shape and observed information about both stimulus properties in the same brain regions. First, Kaiser *et al.*⁶ used cross-decoding searchlight to locate information about shape matched images of clothing (gloves and shirts) and body parts (hands and torsos). Cross-decoding revealed information about shape in EVC but not category, with overlapping clusters for both in LOC. Second, Bracci and Op de Beeck⁷ used a more complex design of 6 object categories and 9 shape types, and found information for both stimulus properties was present in multiple sub-regions of LOC and VTC. In contrast to these studies, we only observed an overlap in object category and shape information in LOC and VTC when carrying out a searchlight analysis with standard cross-validation. When using cross-decoding, information about shape was confined to EVC, while information about animacy was still found in LOC and to a lesser extent VTC; though a small RT-distance effect was still observed in LOC for the bar-like objects. In our stimulus design high and low-level shape are less conflated for the bar-like stimuli, which might account for why there is still a shape effect in LOC. More generally, an important difference between the present study and these previous findings is that low- and high-level shape were not differentiated in their stimulus designs.

The results for the indoor-outdoor task are difficult to interpret, but suggest that representations related to object category and shape might first be recruited, and the timing in the activation of these representations then impact performance on the indoor-outdoor task. It is also notable that both clusters for the indoor category, pets and tools, showed significant effects in LOC. This may be because pets may produce a more decodable animacy response than insects³⁴, and there is evidence that sub-regions of LOC code specifically for tools^{35,36}. Importantly, the size of the significant correlations in the indoor-outdoor task is much smaller than some of the correlations in the two other tasks, in absolute terms as well as relative to the reliability (explainable variance).

In summary, previous studies have used neural distance to predict observer behavior on visual categorization tasks for animacy. We built on these findings using a stimulus design that orthogonalized object category and shape, while also introducing a number of methodological innovations to NDBA⁸. In line with previous results, we observed a robust effect for animacy in LOC driven by the animate stimuli, as well as a shape effect in EVC, and

a smaller crisscross effect using a more contrived category structure that crisscrossed the two dimensions. Taken together these results suggest there is potential to expand the neural distance-to-bound approach to other divisions beyond animacy and object category per se¹⁷.

Data Availability

Raw and processed neural and behavioral data are available on the Open Science Framework (<https://osf.io/3d-pqe/>). Analysis files required to fully replicate all aspects of the study are available from the corresponding author on reasonable request.

References

1. Grill-Spector, K. & Weiner, K. S. The functional architecture of the ventral temporal cortex and its role in categorization. *Nat. Rev. Neurosci.* **15**, 536–548 (2014).
2. Andrews, T. J., Watson, D. M., Rice, G. E. & Hartley, T. Low-level properties of natural images predict topographic patterns of neural response in the ventral visual pathway. *J. Vis.* **15**, 1–12 (2015).
3. Rice, G. E., Watson, D. M., Hartley, T. & Andrews, T. J. Low-level image properties of visual objects predict patterns of neural response across category-selective regions of the ventral visual pathway. *J. Neurosci.* **34**, 8837–8844 (2014).
4. Coggan, D. D., Liu, W., Baker, D. H. & Andrews, T. J. Category-selective patterns of neural response in the ventral visual pathway in the absence of categorical information. *Neuroimage* **135**, 107–114 (2016).
5. Proklova, D., Kaiser, D. & Peelen, M. V. Disentangling representations of object shape and object category in human visual cortex: the animate–inanimate distinction. *J. Cogn. Neurosci.* **28**, 680–692 (2016).
6. Kaiser, D., Azzalini, D. C. & Peelen, M. V. Shape-independent object category responses revealed by MEG and fMRI decoding. *J. Neurophysiol.* **115**, 2246–2250 (2016).
7. Bracci, S. & Op de Beeck, H. Dissociations and associations between shape and category representations in the two visual pathways. *J. Neurosci.* **36**, 432–444 (2016).
8. Bracci, S., Ritchie, J. B. & Op de Beeck, H. On the partnership between neural representations of object categories and visual features in the ventral visual pathway. *Neuropsychologia* **105**, 153–164 (2017).
9. Ritchie, J. B., Kaplan, D. M. & Klein, C. Decoding the Brain: Neural Representation and the Limits of Multivariate Pattern Analysis in Cognitive Neuroscience. *Br. J. Philos. Sci.* **70**, 581–607 (2019).
10. de-Wit, L., Alexander, D., Ekroll, V. & Wagemans, J. Is neuroimaging measuring information in the brain? *Psychon. Bull. Rev.* **23**, 1415–1428 (2016).
11. Tong, F. & Pratte, M. S. Decoding patterns of human brain activity. *Annu. Rev. Psychol.* **63**, 483–509 (2012).
12. Ashby, F. G. & Maddox, W. T. A response time theory of separability and integrality in speeded classification. *J. Math. Psychol.* **38**, 423–466 (1994).
13. Pike, R. Response latency models for signal detection. *Psychol. Rev.* **80**, 53 (1973).
14. Ritchie, J. B. & Carlson, T. A. Neural decoding and ‘inner’ psychophysics: a distance-to-bound approach for linking mind, brain, and behavior. *Front. Neurosci.* **10**, 1–8 (2016).
15. Carlson, T. A., Ritchie, J. B., Kriegeskorte, N., Durvasula, S. & Ma, J. Reaction time for object categorization is predicted by representational distance. *J. Cogn. Neurosci.* **26**, 132–142 (2014).
16. Ritchie, J. B., Tovar, D. A. & Carlson, T. A. Emerging object representations in the visual system predict reaction times for categorization. *PLOS Comput. Biol.* **11**, e1004316 (2015).
17. Grootswagers, T., Cichy, R. M. & Carlson, T. A. Finding decodable information that can be read out in behaviour. *Neuroimage* **179**, 252–262 (2018).
18. Grootswagers, T., Ritchie, J. B., Wardle, S. G., Heathcote, A. & Carlson, T. A. Asymmetric compression of representational space for object animacy categorization under degraded viewing conditions. *J. Cogn. Neurosci.* **29**, 1995–2010 (2017).
19. Brainard, D. H. The psychophysics toolbox. *Spat. Vis.* **10**, 433–436 (1997).
20. Oliva, A. & Torralba, A. Modeling the shape of the scene: a holistic representation of the spatial envelope. *Int. J. Comput. Vis.* **42**, 145–175 (2001).
21. Op de Beeck, H. P. Probing the mysterious underpinnings of multi-voxel fMRI analyses. *Neuroimage* **50**, 567–571 (2010).
22. Oosterhof, N. N., Connolly, A. C. & Haxby, J. V. CoSMoMVPA: multi-modal multivariate pattern analysis of neuroimaging data in Matlab/GNU Octave. *Front. Neuroinform.* **10**, 1–27 (2016).
23. Mur, M., Bandettini, P. A. & Kriegeskorte, N. Revealing representational content with pattern-information fMRI—an introductory guide. *Soc. Cogn. Affect. Neurosci.* 1–9 (2009).
24. Vandierendonck, A. A comparison of methods to combine speed and accuracy measures of performance: a rejoinder on the binning procedure. *Behav. Res. Methods* **49**, 653–673 (2017).
25. DiCarlo, J. J. & Johnson, K. O. Velocity invariance of receptive field structure in somatosensory cortical area 3b of the alert monkey. *J. Neurosci.* **19**, 401–419 (1999).
26. Op De Beeck, H. P., Deutsch, J. A., Vanduffel, W., Kanwisher, N. G. & DiCarlo, J. J. A stable topography of selectivity for unfamiliar shape classes in monkey inferior temporal cortex. *Cereb. Cortex* **18**, 1676–1694 (2008).
27. Cronbach, L. J. *Essentials of Psychological Testing*. (Oxford, England: Harper, 1949).
28. Cichy, R. M., Chen, Y. & Haynes, J.-D. Encoding the identity and location of objects in human LOC. *Neuroimage* **54**, 2297–2307 (2011).
29. Connolly, A. C. *et al.* The representation of biological classes in the human brain. *J. Neurosci.* **32**, 2608–2618 (2012).
30. Eger, E., Ashburner, J., Haynes, J.-D., Dolan, R. J. & Rees, G. fMRI activity patterns in human LOC carry information about object exemplars within category. *J. Cogn. Neurosci.* **20**, 356–370 (2008).
31. Connolly, A. C. *et al.* How the human brain represents perceived dangerousness or ‘predacity’ of animals. *J. Neurosci.* **36**, 5373–5384 (2016).
32. Ritchie, J. B., Bracci, S. & Op de Beeck, H. Avoiding illusory effects in representational similarity analysis: what (not) to do with the diagonal. *Neuroimage* **148**, 197–200 (2017).
33. Walther, A. *et al.* Reliability of dissimilarity measures for multi-voxel pattern analysis. *Neuroimage* **137**, 188–200 (2016).
34. Sha, L. *et al.* The animacy continuum in the human ventral vision pathway. *J. Cogn. Neurosci.* **27**, 665–678 (2015).
35. Bracci, S., Cavina-Pratesi, C., Ietswaart, M., Caramazza, A. & Peelen, M. V. Closely overlapping responses to tools and hands in left lateral occipitotemporal cortex. *J. Neurophysiol.* **107**, 1443–1456 (2011).
36. Bracci, S., Cavina-Pratesi, C., Connolly, J. D. & Ietswaart, M. Representational content of occipitotemporal and parietal tool areas. *Neuropsychologia* **84**, 81–88 (2016).
37. Van Essen, D. C. *et al.* An integrated software suite for surface-based analyses of cerebral cortex. *J Am Med Inform.* **8**, 443–459 (2001).

Acknowledgements

This work was supported by the European Research Council (ERC-2011-StG-284101), a federal research action (IUAP-P7/11), a Hercules grant ZW11_10, and a KU Leuven research council grant (C14/16/031) to H.O. This project has received funding from the FWO and European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 665501, via a FWO [PEGASUS]² Marie Skłodowska-Curie fellowship (12T9217N) to J.B.R.

Author Contributions

Both authors jointly designed the experiment. J.B.R. collected and analyzed the data, and wrote the main manuscript and prepared the figures. Both authors reviewed the manuscript.

Additional Information

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019