Yellow strawberries and red bananas: The influence of object-colour knowledge on emerging object representations in the brain

Lina Teichmann[1,2], Genevieve L. Quek[5], Amanda K. Robinson[1,2,3], Tijl Grootswagers[1,2,3], Thomas A. Carlson[2,3], Anina N. Rich*[1,2,4]

1| Perception in Action Research Centre & Department of Cognitive Science, Macquarie University, Sydney, Australia
2| ARC Centre of Excellence in Cognition & its Disorders, Macquarie University, Sydney, Australia
3| School of Psychology, University of Sydney, Australia
4| Centre for Elite Performance, Training and Expertise, Macquarie University, Sydney, Australia
5| Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands

*Corresponding author's email: anina.rich@mq.edu.au

Keywords: Object-colour knowledge, feature binding, MEG, MVPA, decoding

ACKNOWLEDGEMENTS:
This study was conducted at the KIT-Macquarie Brain Research (MEG) Laboratory. The research was supported by the Australian Research Council (ARC) Centre of Excellence in Cognition and its Disorders, International Macquarie University Research Training Program Scholarships to LT & TG, an ARC Future Fellowship (FT120100816) and an ARC Discovery project (DP160101300) to TC. ANR has funding from the ARC (DP12102835 and DP170101840). The authors acknowledge the University of Sydney HPC service for providing High Performance Computing resources.

Abstract

The ability to rapidly and accurately recognise complex objects is a crucial function of the human visual system. Successful object recognition requires binding incoming visual features such as colour and form into specific neural representations that can be compared to our pre-existing knowledge about the world. For some objects, typical colour is a central feature for recognition; for example, a banana is typically yellow. Here, we examine the timecourse over which features such as colour and form are bound together by using multivariate pattern analyses of time-resolved neuroimaging (magnetoencephalography) data. Consistent with a traditional hierarchical view, we show that single object features are processed before the features are bound into a coherent object that can be compared with existing, conceptual object representations. Our data also suggest that colour processing is be affected by the conjunction of object and colour. These results provide new insights into the interaction between our knowledge about the world and incoming visual information.

Introduction

Successful object recognition depends critically on comparing incoming perceptual info with existing internal representations (Albright, 2012; Clarke & Tyler, 2015). A central feature of many objects is colour, which can be a highly informative cue about an object's identity. For example, when we see a small oval fruit that is yellow in colour, we know that we are looking at a lemon and not a lime. Although a lot is known about colour perception itself, we know comparatively less about how object-colour knowledge interacts with colour perception and object processing. Here, we apply multivariate pattern analyses (MVPA) to Magnetoencephalography (MEG) data to examine how the activation of object-colour knowledge unfolds over time.

There is substantial behavioural evidence that our existing knowledge about an object's typical colour interacts with processing perceptual object features. From the behavioural literature, we know that representations of canonically-coloured objects inherently include colour as a strong defining feature, such that conflicting colour information (e.g., a red banana) slows recognition (Nagai & Yokosawa, 2003; Tanaka & Presnell, 1999, for a meta-analysis see Bramão, Reis, Petersson, & Faísca, 2011). Neuroimaging and neural stimulation experiments suggests that this *binding* of incoming perceptual information and object-colour knowledge takes place in the anterior temporal lobe (ATL) (Chiou, Sowman, Etchell, & Rich, 2014; Coutanche & Thompson-Schill, 2014; Pobric, Jefferies, & Lambon Ralph, 2010). In one study, for example, brain activation patterns evoked by recalling a known object's colour and its shape could be distinguished in brain areas that have consistently been associated with those features, namely V4 and lateral occipital cortex (LOC) respectively (Coutanche & Thompson-Schill, 2014). In contrast, recalling an object's particular conjunction of colour *and* shape, that is, a 'bound' representation, could only be distinguished in the anterior temporal lobe (ATL). Similarly, results from patient work (Patterson, Nestor, & Rogers, 2007) and transcranial magnetic stimulation studies (Chiou et al., 2014; Pobric et al., 2010) point towards the ATL as a hub for conceptual knowledge (for a recent review see Ralph, Jefferies, Patterson, & Rogers, 2017). While these results suggest that the ATL carries conceptual information, it is unclear how conceptual-level processing interacts dynamically with perception.

Time-resolved data, such Electroencephalography (EEG) or MEG data, can give an understanding of the stage of processing at which incoming perceptual information is influenced by stored object-knowledge. Previous EEG studies have examined the temporal dynamics of object-colour knowledge as an index of the integration of incoming visual information and prior knowledge (Lloyd-Jones, Roberts, Leek, Fouquet, & Truchanowicz, 2012; Lu et al., 2010; Proverbio, Burco, del Zotto, & Zani, 2004). For example, Lloyd-Jones et al. (2012) showed participants images of everyday objects coloured correctly (e.g., a yellow banana) or incorrectly (e.g., a purple banana) while recording EEG data. Their results show that shape information modulated the neural responses at around 170ms (i.e., component N1), the combination of shape and colour affected the signal at 225ms (i.e., component P2), and the typicality of object-colour pairing modulated components approximately 225 and 350ms after stimulus onset (i.e., P2 and P3). These findings suggest that shape information activates typical object-colour associations and that bound colour *and* shape features are processed later than shape or colour alone. This suggests that the initial stages of object recognition may be purely based on shape, with the interactions with object-colour knowledge coming into play at a much later stage, perhaps as late as during response selection.

In these previous studies, the focus was on evoked components, which cannot tell us about the type of information that is contained in the neural signal. In the present study, we examine the temporal aspects underlying object-colour processing using time-resolved multivariate analyses of MEG data, which allows us to explore when particular types of information (e.g., shape, colour, congruency) influence neural activity. This provides a unique insight into the temporal dynamics of object-knowledge and object-feature binding by showing how existing knowledge about an object's typical colour affects perceptual processing of that object's features. We presented participants with coloured objects that were either congruent (e.g., yellow banana) or incongruent (e.g., red banana). Using machine learning algorithms, we determined the timepoint at which neural activity differed between congruently and incongruently coloured objects, which reflects the time by which binding of colour and shape must have occurred. By further contrasting the neural responses evoked by congruent and incongruently coloured objects with those evoked by objects without colour (e.g., greyscale banana) and colours without familiar objects (colours overlaid on abstract shapes), we also examine whether existing knowledge about an object's typical colour influences perceptual processing of those features. Overall, our findings elucidate the timecourse of interactions between incoming visual information and prior knowledge in the brain, demonstrating the importance of what we know in determining what we see.

Methods

Participants. 20 healthy volunteers (11 female, mean age = 28.9 years, SD = 6.9 years, 1 left-handed) participated in the study. All participants reported accurate colour-vision and had normal or corrected-to-normal visual acuity. Participants gave informed consent before the experiment started and were financially compensated. The study was approved by the Macquarie University Human Research Ethics Committee.

Stimuli. We identified five real world objects that previous studies have shown to be strongly associated with each of four different colours (red, green, orange and yellow; see Figure 1) (Bannert & Bartels, 2013; Joseph, 1997; Lloyd-Jones et al., 2012; Naor-Raz, Tarr, & Kersten, 2003; Tanaka & Presnell, 1999; Therriault, Yaxley, & Zwaan, 2009). Each colour category had one manmade object (e.g., fire hydrant), one living object (e.g., ladybird), and three fruits or vegetables (e.g., strawberry, tomato, cherry). We sourced two exemplar images for each object class, resulting in 10 images for each colour, 40 individual images in total. We then created incongruently coloured objects by swapping the colours (e.g., green strawberry, orange broccoli). For both congruent and incongruent stimuli, we did not use the native colours from the images themselves, but instead overlayed pre-specified hues on desaturated (greyscale) images that were equated for luminance using the SHINE toolbox (Willenbockel et al., 2010). This ensured that congruent and incongruent stimuli were matched in the way the texture and shape of the object interacted with the colour overlay. A greyscale image overlayed with its canonically associated colour (e.g., yellow hue applied to greyscale banana) resulted in a congruent object; a greyscale image overlayed with a colour different from its canonically associated colour (e.g., red hue applied to greyscale banana) resulted in an incongruent object. Every congruent object exemplar had a single colour-matched incongruent partner. For example, we used a specific shade of red and added it to the grey-scale images of the strawberries to make the congruent strawberries and overlayed it onto the lemons to make the incongruent lemons. We then took a specific shade of yellow and overlayed it on the lemons to make the congruent lemon exemplar, and onto the strawberries to make the incongruent strawberry exemplar. That means, overall, we have the identical shapes and colours in the congruent and the incongruent condition, a factor that is crucial to ensure our results cannot be explained by features other than colour congruency. The only difference between these key conditions is that the colour-shape combination is either typical (congruent) or atypical (incongruent).

This procedure resulted in 40 congruent objects (10 of each colour), and 40 incongruent objects (10 of each colour, Figure 1). We added two additional stimulus types to this set: the full set of 40 greyscale images, and a set of 10 different angular abstract shapes, coloured in each of the four hues for a set of 40 (see Figure 1). As is clear in Figure 1, the colours of the abstract shapes appeared brighter than the colours of the objects, this is because the latter were made by overlaying hue on greyscale, whereas the shapes were simply coloured. As our principle goal was to ensure that the congruent objects appeared to have their typical colouring, we did not match the overall luminance of the coloured stimuli. For example, if we equated the red of a cherry with the yellow of a lemon, neither object would look typically coloured. Thus, each specific colour pair is not equated for luminance; however we have the same colours across different conditions, which ensures this cannot form a clue for the classification algorithm in distinguishing our categories.

All stimuli were presented at a distance of 114cm and image size varied randomly from trial to trial by 2 degrees visual angle resulting in the visual angle of ~4.3 – 6.3 degrees. This added visual variability to reduce low-level featural differences not related to colour between images.

Procedure. Before entering the magnetically shielded room for MEG recordings, an elastic cap with five marker coils was placed on the participant's head. We recorded head shape with a digitiser pen and used these marker coils to measure the head position within the magnetically shielded room at the start of the experiment, half way through and at the end.

In the main task (Figure 1B), participants completed eight blocks of 800 trials each. Each individual stimulus appeared 40 times over the course of the experiment. Each stimulus was presented centrally for 450ms with a black fixation dot on top of it. To keep participants attentive, after every 80 trials, a target image was presented until a response was given indicating whether this stimulus had appeared in the immediately previous block of trials or not (50% present vs absent). The different conditions (congruent, incongruent, grey-scale, abstract shape) were randomly intermingled throughout each block, and the target was randomly selected each time. On average, participants performed with 90% (SD=5.4%) accuracy.

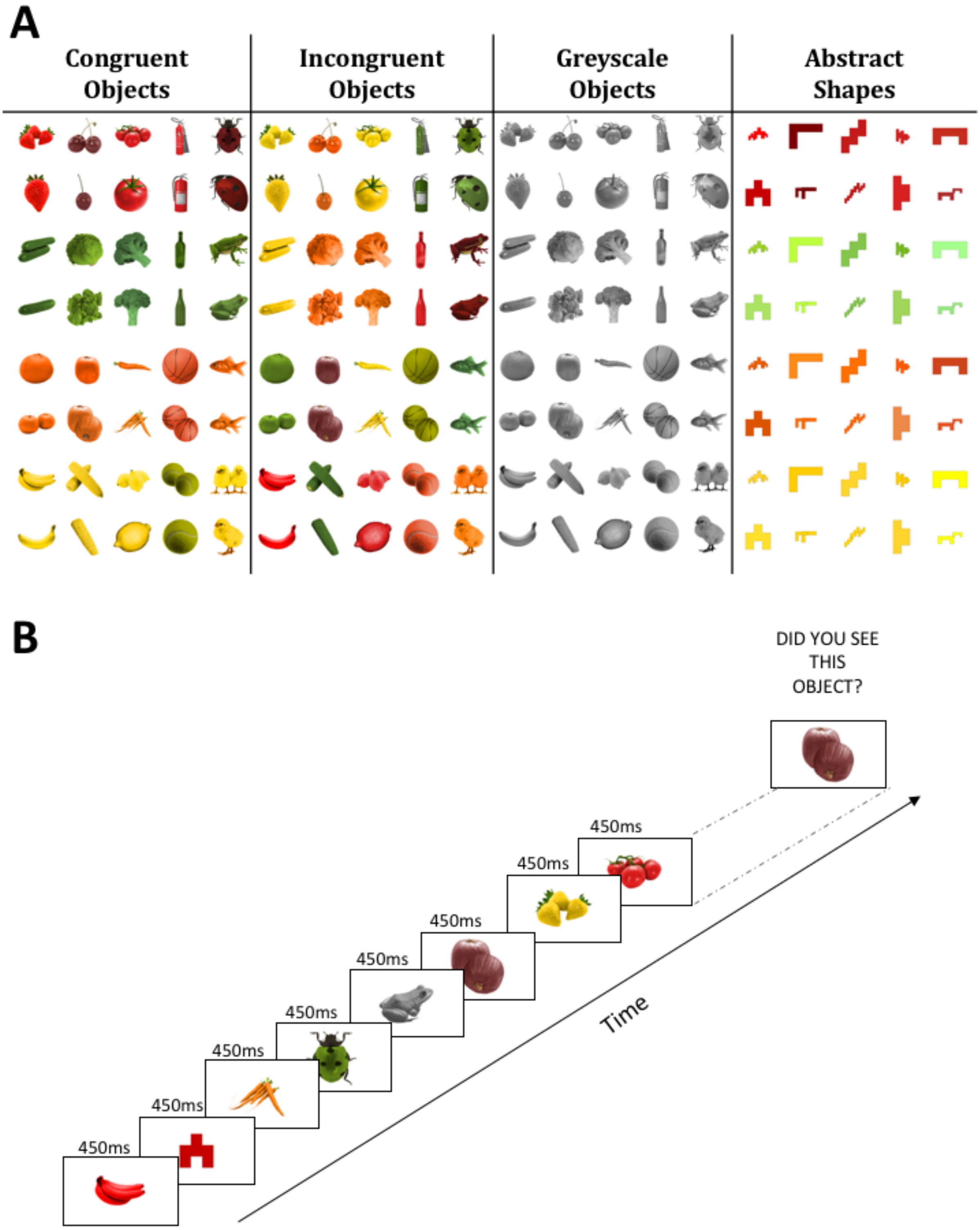

Figure 1. (A) shows all stimuli used in this experiment. The same objects were used in the congruent, incongruent, and greyscale conditions. There were two exemplars of each objects. Colours in the congruent and incongruent condition were matched. The abstract shapes were identical across colour categories. (B) shows an example sequence of the main task. Participants viewed each object for 450ms. After each run, one object was displayed and participants had to indicate whether they had seen this object in the previous run or not.

After completing the main blocks, we collected behavioural object-naming data to test for a behavioural congruency effect with our stimuli. On the screen, participants saw each of the objects again (congruent, incongruent or grey-scale) in a random order and were asked to name the objects as quickly as possible. As soon as voice onset was detected, the stimulus disappeared. We marked stimulus-presentation time with a photodiode and recorded voice-onset with a microphone. Seventeen participants completed three blocks of this reaction time task, one participant completed two blocks, and for two participants we could not record any reaction times. Each block contained all congruent, incongruent and grey-scale objects presented once.

Naming reaction times were defined as the difference between stimulus-onset and voice-onset. Trials containing naming errors and microphone errors were not analysed. We calculated the median naming time for each exemplar for each person and then compared the naming times for each of the congruent, incongruent and greyscale conditions.

Apparatus. The neuromagnetic recordings were conducted with a whole-head axial gradiometer MEG (KIT, Kanazawa, Japan), containing 160 axial gradiometers. We recorded the MEG signal with a 1000Hz frequency. An online low-pass filter of 200Hz and a high-pass filter of 0.03Hz were used. All stimuli were projected on a translucent screen mounted on the ceiling of the magnetically shielded room. Stimuli were presented using MATLAB with Psychtoolbox extension (Brainard, 1997; Brainard & Pelli, 1997; Kleiner et al., 2007). Parallel port triggers and the signal of a photodiode were used to mark the beginning and end of each trial. A Bimanual 4-Button Fiber Optic Response Pad (Current Designs, Philadelphia, USA) was used to record the responses. Head shape recordings were completed with a Polhemus Fastrak digitiser pen (Colchester, USA).

Pre-processing. FieldTrip (Oostenveld, Fries, Maris, & Schoffelen, 2011) was used to pre-process the data. The data were downsampled to 200Hz and then epoched from -100 to

450ms relative to stimulus onset. We did not conduct any further pre-processing steps (filtering, channel selection, trial-averaging etc.) to keep the data in its rawest possible form.

Decoding Analyses. For all our decoding analyses, patterns of brain activity were extracted across MEG sensors at every timepoint, for each participant separately. We used a regularised linear discriminant analysis (LDA) classifier which was trained to distinguish the conditions of interest. We then used independent test data to assess whether the classifier could predict the condition above chance in the new data. We conducted training and testing at every timepoint and tested for significance using random-effects Monte Carlo cluster (TFCE; Smith & Nichols, 2009) statistics, corrected for multiple comparisons using the max statistic across time points (Maris & Oostenveld, 2007). Note that our aim was not to achieve the highest possible decoding accuracy (i.e., "classification for prediction", Hebart & Baker, 2017), but rather to test whether the classifier could predict the conditions above chance at any of the timepoints (i.e., "classification for interpretation", Hebart & Baker, 2017). Therefore, we followed a minimal preprocessing pipeline and performed our analyses on a single-trial basis. Classification accuracy above chance indicates that the MEG data contains information that is different for the categories. We used the CoSMoMVPA toolbox (Oosterhof, Connolly, & Haxby, 2016) to conduct all our analyses.

We ran three decoding analyses to examine how the typicality of object-colour combinations influences colour and shape processing over time. By examining the timecourse of object-feature binding, these analyses allow us to track the interaction between object-colour knowledge and object representations in the brain. First, we tested whether activation patterns evoked by congruently coloured objects (e.g., red strawberry) differ from activation patterns evoked by incongruently coloured objects (e.g., yellow strawberry). Any differential response that depends on whether a colour is typical or atypical for an object (a congruency effect) requires the perceived shape and colour to be bound and compared to a conceptual object representation activated from memory. We trained the classifier on all congruent and incongruent trials *except* for trials corresponding to one pair of matched exemplars (e.g., all instances of congruent and incongruent strawberries and congruent and incongruent bananas). We then tested the classifier using only the left-out exemplar pairs. We repeated this process until each matched exemplar pair had been left out (i.e., used as test data) once. Leaving an exemplar pair out ensures that there are identical shapes and colours for both classes (congruent and incongruent) in both the training and the testing set, and that the stimuli of the test set have different shape characteristics than any of the training objects. As such, the only distinguishing feature between the conditions is the *conjunction* of shape and colour features, which defines congruency. This allows us to compare directly whether (and at which timepoint) object-colour knowledge interacts with stored object representations.

In a second decoding analysis, we examined whether the conjunction of object and colour influenced colour perception itself. Perceiving a strongly associated shape in the context of viewing a certain colour might lead to a more stable representation of that colour in the MEG signal. For example, if we see a yellow banana, the banana shape may facilitate a stable representation of the colour yellow earlier than if we see a yellow strawberry. To assess this possibility, we trained the classifier to distinguish between the surface colours of the abstract shapes (i.e., red, orange, yellow, green, chance: 25%). We then tested how well the classifier could predict the colour of the congruent and incongruent objects. Training the classifier on the same abstract shapes across colour categories makes it impossible that a certain shape-colour combination drives an effect, as the distinguishing feature between the abstract shapes is colour. This analysis allows us to compare whether the typicality of colour-form combinations has an effect on colour processing.

Third, we tested whether the conjunction of object and colour has an effect on object decoding. If object-colour influences early perceptual processes, we might see a facilitation for decoding objects when they are coloured congruently or interference when the objects are coloured incongruently. We used the greyscale object trials to train the classifier to distinguish between all of the objects. The stimulus set contained two exemplars of each item (e.g., strawberry 1 and strawberry 2). We used different exemplars for the training and testing set to minimise the effects of low-level visual features, however, given that there are major differences in object shapes and edges we can still expect to see strong differences between the objects. The classifier was trained on one exemplar of all of the greyscale trials. We then tested the classifier's performance on the congruent and incongruent object trials using the exemplars the classifier did not train on. We then swapped the exemplars used for training and testing set until every combination had been used in the testing set. Essentially, this classifier is trained to predict which object was presented to the participant (e.g., was it a strawberry or a frog?) and we are testing whether there is a difference depending on whether the object is congruently or incongruently coloured.

In addition to our main analyses, we also conducted additional decoding analyses to replicate and extend an earlier study testing the timecourse of colour processing (Teichmann, Grootswagers, Carlson, & Rich, 2019). We tested whether colour representations accessed via perception (i.e., coloured abstract shapes) and via associations (i.e., greyscale objects associated with a colour) evoke similar neural patterns. The results from these additional analyses are summarised in the supplementary materials.

Statistical Tests. In all our analyses, we used random effects Monte-Carlo cluster statistic using Threshold Free Cluster Enhancement (TFCE, Smith & Nichols, 2009) as implemented in the CoSMoMVPA toolbox to see whether the classifier could predict the decisions above chance. The TFCE statistic represents the support from neighbouring time points, thus allowing for detection of sharp peaks and sustained small effects over time. We used a permutation test, swapping labels of complete trials, and re-ran the decoding analysis on the data with the shuffled labels 100 times per participant to create subject-level null-distributions. We then used Monte-Carlo sampling to create a group-level null-distribution consisting of 10,000 shuffled label permutations for the time-resolved decoding, and 1000 for the time-generalisation analyses (to reduce computation time). The null distributions were then transformed into TFCE statistics. To correct for multiple comparisons, the *maximum* TFCE values across time in each of the null distributions was selected. We then transformed the true decoding values to TFCE statistics. To assess whether the true TFCE value at each timepoint is significantly above chance, we compared it to the 95th percentile of the corrected null distribution. Selecting the maximum TFCE value provides a conservative threshold for determining whether the observed decoding accuracy is above chance, corrected for multiple comparisons.

Results

Behavioural results. We first present the behavioural data to confirm that our stimuli induce a congruency effect on object naming times. All incorrect responses and microphone errors were excluded from analysis (on average across participants: 10.1%). We then calculated the median reaction time for naming each stimulus. If a participant named a specific stimulus incorrectly across trials (e.g., incongruently coloured strawberry was always named incorrectly), we removed this stimulus completely to ensure that the reaction times in one condition were not skewed (on average this occurred in 5.4% of cases). Participants were faster to name the congruently coloured (702ms) than incongruently coloured (750ms) objects ($t(17) = 4.06$, $p < .001$; 95% CI [22.9, 72.7]). This suggests that the objects we used here do indeed have associations with specific canonical colours, and we replicate the effect of colour congruency on recognition of these objects (Bannert & Bartels, 2013; Joseph, 1997; Lloyd-Jones et al., 2012; Naor-Raz et al., 2003; Tanaka & Presnell, 1999; Therriault et al., 2009).

In the main task, participants were asked to indicate every 80 trials whether they had seen a certain target object or not. The aim of this task was to keep participants motivated and attentive throughout the training session. On average, participants reported whether the targets were present or absent with 90% accuracy (SD = 5%, range: 81.25% - 100%).

MEG decoding results. The aim of our decoding analyses was to examine the interaction between object-colour knowledge and object representations. First, we tested for a difference in the brain activation pattern for congruently and incongruently coloured objects. The results show distinct patterns of neural activity for congruent compared to incongruent objects in a cluster stretching from 265 to 330ms after stimulus onset, demonstrating that brain activity is modulated by colour congruency in this time window (Figure 2B). Thus, binding of colour and form must have occurred by ~265ms.

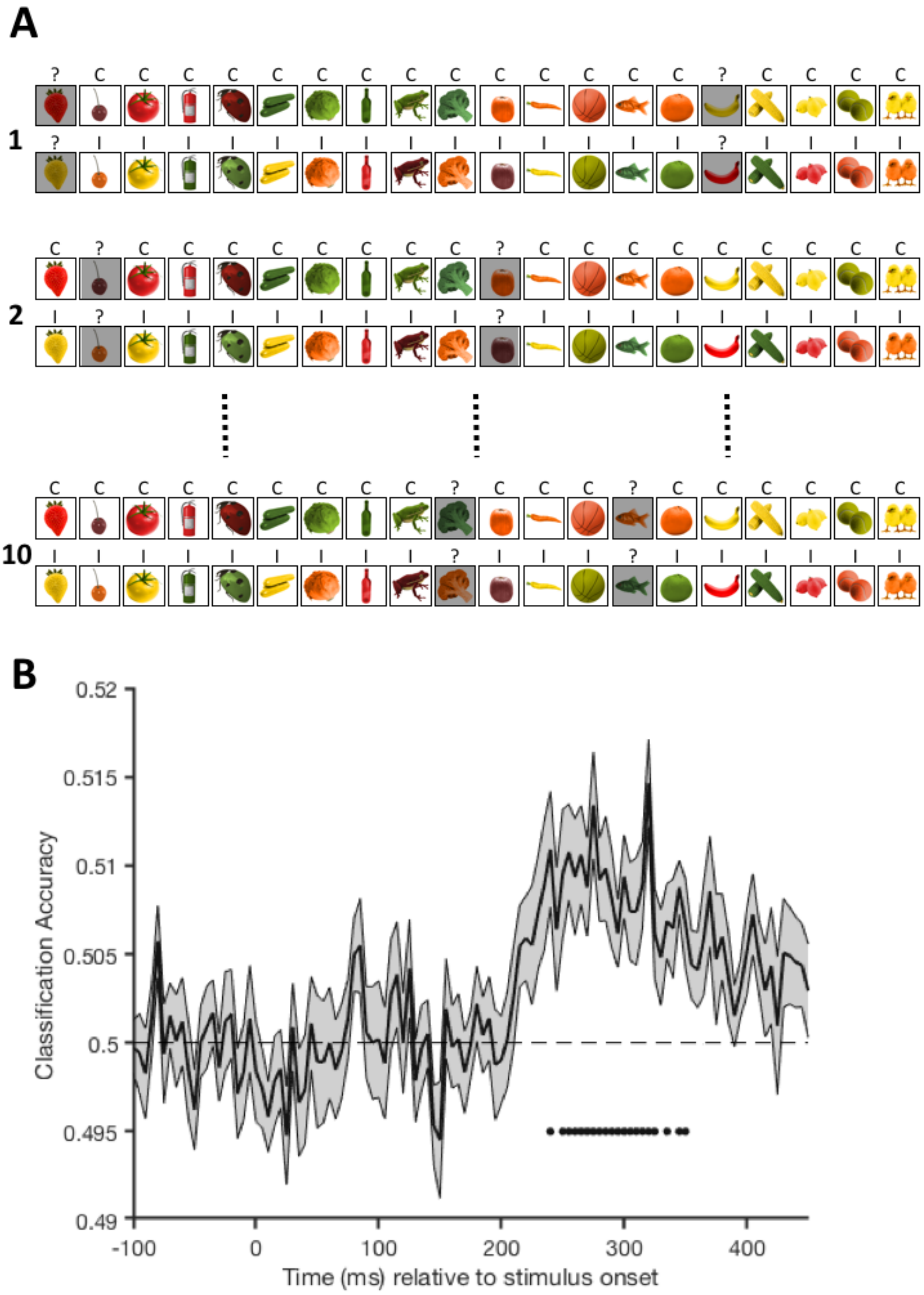

Figure 2. Cross-validation and results of the congruency analysis contrasting trials from the congruent and incongruent conditions. (A) shows the leave-one-matched-exemplar-out cross validation approach used for the congruency decoding analysis. The classifier was trained on the trials shown in clear squares and then tested on the trials shaded grey (actual stimuli were all clear), ensuring that the classifier is not tested on the exemplars that it trained on. This limits the effect features other than congruency can have on classifier performance. (B) shows the classification accuracy over time. Shading represents the standard error across participants. Black dashed line represents chance level (50% - congruent versus incongruent). Black dots highlight significant timepoints, corrected for multiple comparisons.

To assess whether congruency influences colour perception, we trained a classifier to distinguish between the colours in the abstract shape condition and then tested it on the congruent and incongruent trials separately (see supplementary materials for further colour decoding analyses). Colour can be successfully classified in a cluster stretching from 75 to 125ms for the congruent condition and in a cluster stretching from 75 to 185ms for the incongruent trials (Figure 3A). These results suggest there is a qualitative difference between the way colour information is processed depending on the congruency of the image. To assess how these signals evolves over time, we used time-generalisation matrices (Figure 3B and 3C). Colour category was decodable from both conditions early on (~70ms) but in the incongruent condition, the activation associated with colour seems to be sustained longer (Figure 3C) than for the congruent condition (Figure 3B). This suggests that colour signals are prolonged when object-colour combinations are unusual relative to when they are typical.

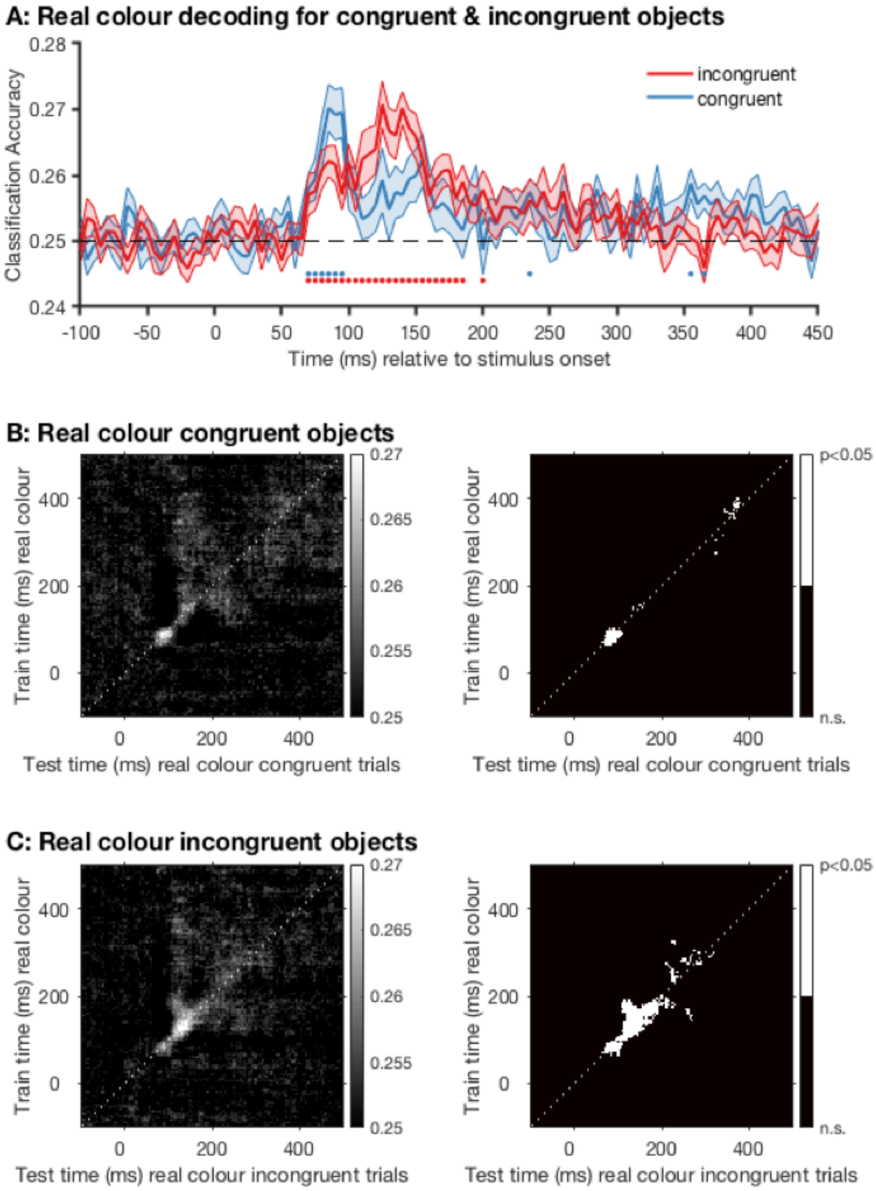

Figure 3. Results of the colour decoding analysis for the congruent and incongruent trials. Here, the classifier was trained to distinguish the colour of all abstract shape trials and tested on the congruent and incongruent trials separately. (A) shows the classification accuracy over time for this analysis. Red indicates the classifier's performance when it was tested on incongruent trials and blue indicates the classifier's performance when it was tested on congruent trials. Shading represents the standard error. Black dashed line indicates chance level (25% - red versus green versus orange versus yellow). Red (incongruent) and blue (congruent) dots highlight significant timepoints, corrected for multiple comparisons. (B) and (C) show the results of the same analysis across all possible training and testing timepoint combinations. These time-time matrices allow us to examine how the signal for the congruent colours (B) and incongruent colours (C) evolves over time. The plots on the left show the classification accuracy at every timepoint combination with lighter pixels reflecting higher decoding accuracies. The plots on the right show clusters where decoding is significantly above chance, corrected for multiple comparisons.

The goal of the third analysis was to examine whether shape representations are affected by colour congruency. We trained a classifier to distinguish between trials in which the participant saw one of the exemplars of each of the twenty objects in greyscale (e.g., greyscale strawberry 1, greyscale cherry 1, etc.). We then tested at which timepoint the classifier could successfully cross-generalise to the other exemplar of that object in the congruent and incongruent condition separately. Note that although the exact images are unique, there are shared shape characteristics between exemplars (e.g., the two frog exemplars share some shape aspects despite being different postures) which can be expected to drive the classifier's performance. The results show the neural data has differential information about the object in a cluster stretching from 80 to 450ms for the congruent test set and from 90 to 450ms for the incongruent test set (Figure 4A). These results show that we can decode the object category early on, consistent with the classifier being driven by low-level visual features such as shape or texture. The timecourse for congruent and incongruent exemplar decoding is very similar, suggesting that colour congruency does not affect the initial stages of object processing. On error trials, we can see whether the classifier makes the same mistakes when decoding congruent and incongruent stimuli. Figure 4B shows these errors, averaged across timepoints when the decoding was significant. Figure 4C shows a high correlation between the resulting confusion matrices. This demonstrates that the particular classification errors in congruent and incongruent conditions are highly similar, reflecting shape effects rather than colour congruency effects. Thus, we have no evidence that early stages of object processing are affected by colour congruency.

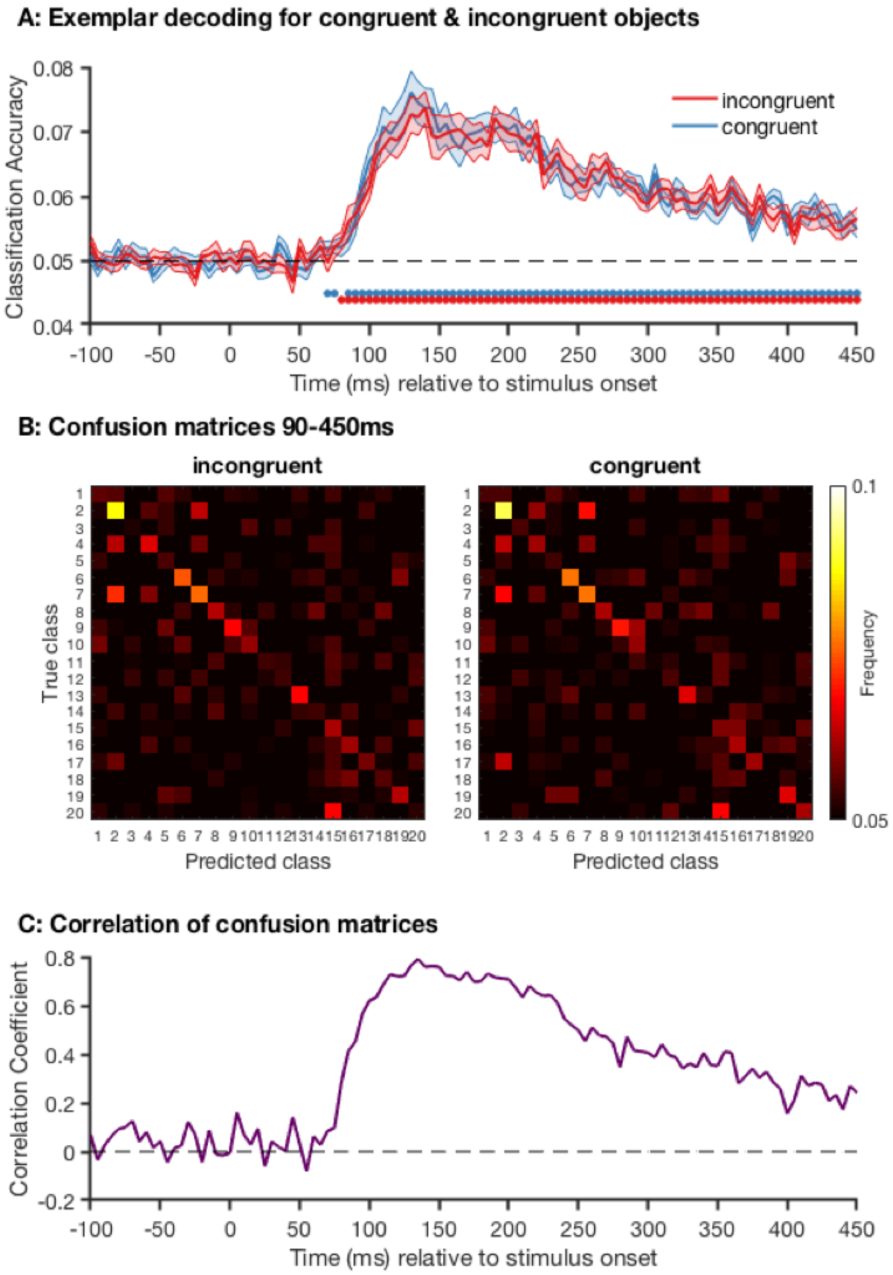

Figure 4. Results of the object exemplar decoding analysis. The classifier was trained to distinguish between all object categories in the greyscale object condition. We used one exemplar of each class for the training and the other exemplar for testing the classifier. Testing was done for the congruent and incongruent trials separately. (A) shows the classification accuracy over time for the exemplar decoding analysis. The red line shows the classifier accuracy when testing the classifier on incongruent trials and blue shows the classifier accuracy when testing the classifier on congruent trials. Shading represents the standard error. Black dashed line represents chance level (5% - 20 different object categories). Red (incongruent) and blue (congruent) dots highlight significant timepoints (p<0.05), corrected for multiple comparisons. (B) shows the confusion matrices for the incongruent (left) and congruent (right) condition. The confusion matrices were averaged across time (90-450ms). The true object categories are plotted on the y-axis and the predicted categories on the x-axis. The colour highlights the frequency of a specific cateogry being predicted in response to a

given category. The diagonal shows accurate classifier prediction (match between predicted and true category). (C) shows the correlation of the two confusion matrices over time.

Discussion

A crucial question in object recognition is how incoming visual information interacts with stored object concepts to create meaningful vision under varying situations. The aims of the current study were to examine the temporal dynamics of object-colour knowledge and to test whether activating object-colour knowledge influences early stages of colour and object processing. Our data provide three major insights: First, congruently and incongruently coloured objects evoke a different neural representation after ~260ms suggesting that by this time, visual object features are bound into a coherent representation and compared to stored object representations. Second, compared with the latency at which congruency decoding is possible, the congruency of the object-colour binding does seem to affect colour processing earlier in the signal. This indicates that there is some information about the "correctness" of an object's colour even in the early stages of processing. Third, we find no evidence that the congruency of object-colour binding affects shape processing, suggesting behavioural congruency effects are due to conflict at a later stage in processing.

Here, we use colour congruency as an index to assess when prior knowledge is integrated with bound object features. When comparing brain activation patterns of the same objects presented in different colours, there was a decodable difference between congruent and incongruent conditions from 265ms onwards suggesting that a stored object representation containing information about the typical colour of an object must have been activated at that stage. Prior to this time, the signal is primarily driven by early perceptual features such as colour and shape, which were matched for the congruent and incongruent conditions (same objects, same colours, only the binding of colour and shape differed). Thus, our data illustrate the classic hierarchy of vision with single features being processed first and the conjunction of colour and shape occurring at a later stage. These timecourse data do not speak to which brain areas are involved in the integration of colour and shape information, which has already been explored by the fMRI literature, instead our congruency analysis shows the relative timecourse at which different features are bound together and a meaningful object representation emerges. These results are consistent with previous work showing that simple colour processes such as registering the intensity of light occurs in early visual areas such as V1 and V2, while more complex colour-related processes such as distinguishing between object surface colours occur in V4 and beyond (Seymour, Williams, & Rich, 2015; Zeki & Marini, 1998). Activating object colour from memory has been shown to involve the ATL (e.g., Coutanche & Thompson-Schill, 2014) and coding for object-colour congruency has involves perirhinal cortex (Price, Bonner, Peelle, & Grossman, 2017). Further support on the involvement of the ATL in the integration of information, such as colour and shape, comes from work on patients with semantic dementia (e.g., Bozeat, Lambon Ralph, Patterson, & Hodges, 2002) and studies on healthy participants using TMS (e.g., Chiou et al., 2014). The latency of congruency decoding in our data may reflect the process of comparing bound perceptual object features with a conceptual template representation of the object in higher-level brain areas such as the ATL.

Our results also show that the “correctness” of an object’s colour has an effect on colour processing. We found colour decoding onset at a similar time (~70ms) for congruently and incongruently coloured objects, however, colours were decodable longer in the incongruent condition than the congruent condition (Figure 3A). Thus, early on colour is processed in a similar way for congruently and incongruently coloured objects, but after the initial, early colour processing, colour information persists only in the incongruent condition. It is possible that this different dynamic in colour processing is driven by certain colours co-occurring more often with other low-level features such as texture, edges or degree of curvature. We presumably learn these regularities from repeated exposure over the lifespan (e.g., Clarke & Tyler, 2015). If there is a conflict between the actual colour and the "most likely" colour, it may lead to the prolonged colour signal which we observed here.

The timecourse of exemplar decoding we observe is consistent with previous studies on object recognition. Here, we found that exemplar identity could be decoded at ~90ms (Figure 4A). Similar latencies have been found in other M/EEG decoding studies (Carlson, Tovar, Alink, & Kriegeskorte, 2013; Cichy, Pantazis, & Oliva, 2014; Contini, Wardle, & Carlson, 2017; Grootswagers, Robinson, & Carlson, 2019; Isik, Meyers, Leibo, & Poggio, 2013) and single unit recordings (e.g., Hung, Kreiman, Poggio, & DiCarlo, 2005). Behavioural data, including the reaction times collected here in our participants, show that colour influences object identification speed (e.g., Bramão, Faísca, Petersson, & Reis, 2010). The neural data, however, did not show an effect of object colour on the classifier's performance when distinguishing the neural activation patterns evoked by different objects. For example, the brain activation pattern in response to a strawberry could be differentiated from the pattern evoked by a lemon, without any influence of the congruency of their colours. This suggests that colour and shape processing affect each other in an asymmetric way: colour representations are influenced by object shape, perhaps due to statistical learning through experience, but shape representations are not influenced by colour. This finding is consistent with previous results (Proverbio et al., 2004) but might seem puzzling because colour congruency has been shown to have a strong effect on object naming (e.g., Chiou et al., 2014; Nagai & Yokosawa, 2003; Tanaka & Presnell, 1999). It seems plausible, however, that the typicality between object and colour combination affects later stages of processing as seen in our congruency analysis, rather than influencing these early stages. For example, the source of behavioural congruency effects may be at the stage of response selection, which would not show up in these early neural signals. More exploration is needed to test this interpretation, but the current data suggest that colour congruency does not have an impact on early stages of shape processing.

Our study demonstrates that object representations are influenced by object-colour knowledge but not at the initial stages of visual processes. Our data also suggest that colour processing is affected by colour congruency, with colour signals being extended for incongruently in comparison to congruently coloured objects. Our findings document the timecourse of the processes suggested by the traditional hierarchical view of vision, in which single object features are processed before the features are bound into a coherent object that can be compared with existing, conceptual object representations. We find that object-colour binding is complete by ~265ms, clearly demonstrating an interaction between our knowledge of the world and incoming information to form our visual perception.

References

Albright, T. D. (2012). On the perception of probable things: Neural substrates of associative memory, imagery, and perception. *Neuron*, *74*(2), 227–245.

Bannert, M. M., & Bartels, A. (2013). Decoding the yellow of a gray banana. *Current Biology*, *23*(22), 2268–2272.

Bozeat, S., Lambon Ralph, M. A., Patterson, K., & Hodges, J. R. (2002). When objects lose their meaning: What happens to their use? *Cognitive, Affective, & Behavioral Neuroscience*, *2*(3), 236–251.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.

Brainard, D. H., & Pelli, D. G. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.

Bramão, I., Faísca, L., Petersson, K. M., & Reis, A. (2010). The influence of surface color information and color knowledge information in object recognition. *The American Journal of Psychology*, *123*(4), 437–446.

Carlson, T. A., Hogendoorn, H., Kanai, R., Mesik, J., & Turret, J. (2011). High temporal resolution decoding of object position and category. *Journal of Vision*, *11*(10), 9–9. https://doi.org/10.1167/11.10.9

Carlson, T. A., Tovar, D. A., Alink, A., & Kriegeskorte, N. (2013). Representational dynamics of object vision: The first 1000 ms. *Journal of Vision*, *13*(10), 1–1.

Chiou, R., Sowman, P. F., Etchell, A. C., & Rich, A. N. (2014). A conceptual lemon: Theta burst stimulation to the left anterior temporal lobe untangles object representation and its canonical color. *Journal of Cognitive Neuroscience*, *26*(5), 1066–1074.

Cichy, R. M., Pantazis, D., & Oliva, A. (2014). Resolving human object recognition in space and time. *Nature Neuroscience*, *17*(3), 455.

Clarke, A., & Tyler, L. K. (2015). Understanding what we see: How we derive meaning from vision. *Trends in Cognitive Sciences*, *19*(11), 677–687.

Contini, E. W., Wardle, S. G., & Carlson, T. A. (2017). Decoding the time-course of object recognition in the human brain: From visual features to categorical decisions. *Neuropsychologia*. https://doi.org/10.1016/j.neuropsychologia.2017.02.013

Coutanche, M. N., & Thompson-Schill, S. L. (2014). Creating concepts from converging features in human cortex. *Cerebral Cortex*, *25*(9), 2584–2593.

Grootswagers, T., Robinson, A. K., & Carlson, T. A. (2019). The representational dynamics of visual objects in rapid serial visual processing streams. *NeuroImage*, *188*, 668–679.

Hebart, M. N., & Baker, C. I. (2018). Deconstructing multivariate decoding for the study of brain function. *Neuroimage*, (180), 4–18.

Hung, C. P., Kreiman, G., Poggio, T., & DiCarlo, J. J. (2005). Fast readout of object identity from macaque inferior temporal cortex. *Science*, *310*(5749), 863–866.

Isik, L., Meyers, E. M., Leibo, J. Z., & Poggio, T. (2013). The dynamics of invariant object recognition in the human visual system. *Journal of Neurophysiology*, *111*(1), 91–102. https://doi.org/10.1152/jn.00394.2013

Joseph, J. E. (1997). Color processing in object verification. *Acta Psychologica*, *97*(1), 95–127.

King, J. R., & Dehaene, S. (2014). Characterizing the dynamics of mental representations: The temporal generalization method. *Trends in Cognitive Sciences*, *18*(4), 203–210.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., Broussard, C., & others. (2007). What's new in Psychtoolbox-3. *Perception*, *36*(14), 1.

Lambon Ralph, M. A., Jefferies, E., Patterson, K., & Rogers, T. T. (2017). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience*, *18*(1), 42.

Lloyd-Jones, T. J., Roberts, M. V., Leek, E. C., Fouquet, N. C., & Truchanowicz, E. G. (2012). The time course of activation of object shape and shape+ colour representations during memory retrieval. *PloS One*, *7*(11), e48550.

Lu, A., Xu, G., Jin, H., Mo, L., Zhang, J., & Zhang, J. X. (2010). Electrophysiological evidence for effects of color knowledge in object recognition. *Neuroscience Letters*, *469*(3), 405–410.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–190.

Nagai, J., & Yokosawa, K. (2003). What regulates the surface color effect in object recognition: Color diagnosticity or category. *Technical Report on Attention and Cognition*, *28*, 1–4.

Naor-Raz, G., Tarr, M. J., & Kersten, D. (2003). Is color an intrinsic property of object representation? *Perception*, *32*(6), 667–680.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, *2011*, 1.

Oosterhof, N. N., Connolly, A. C., & Haxby, J. V. (2016). CoSMoMVPA: Multi-modal multivariate pattern analysis of neuroimaging data in Matlab/GNU Octave. *Frontiers in Neuroinformatics*, *10*. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4956688/

Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience*, *8*(12), 976.

Pobric, G., Jefferies, E., & Lambon Ralph, M. A. (2010). Amodal semantic representations depend on both anterior temporal lobes: Evidence from repetitive transcranial magnetic stimulation. *Neuropsychologia*, *48*(5), 1336–1342.

Price, A. R., Bonner, M. F., Peelle, J. E., & Grossman, M. (2017). Neural coding of fine-grained object knowledge in perirhinal cortex. *BioRxiv*, 194829.

Proverbio, A. M., Burco, F., del Zotto, M., & Zani, A. (2004). Blue piglets? Electrophysiological evidence for the primacy of shape over color in object recognition. *Cognitive Brain Research*, *18*(3), 288–300. https://doi.org/10.1016/j.cogbrainres.2003.10.020

Seymour, K. J., Williams, M. A., & Rich, A. N. (2015). The representation of color across the human visual cortex: Distinguishing chromatic signals contributing to object form versus surface color. *Cerebral Cortex*, *26*(5), 1997–2005.

Smith, S. M., & Nichols, T. E. (2009). Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage*, *44*(1), 83–98.

Tanaka, J. W., & Presnell, L. M. (1999). Color diagnosticity in object recognition. *Perception & Psychophysics*, *61*(6), 1140–1153. https://doi.org/10.3758/BF03207619

Teichmann, L., Grootswagers, T., Carlson, T., & Rich, A. N. (2019). Seeing versus knowing: The temporal dynamics of real and implied colour processing in the human brain. *NeuroImage*.

Therriault, D. J., Yaxley, R. H., & Zwaan, R. A. (2009). The role of color diagnosticity in object recognition and representation. *Cognitive Processing*, *10*(4), 335.

Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., & Tanaka, J. W. (2010). Controlling low-level image properties: The SHINE toolbox. *Behavior Research Methods*, *42*(3), 671–684.

Zeki, S., & Marini, L. (1998). Three cortical stages of colour processing in the human brain. *Brain: A Journal of Neurology*, *121*(9), 1669–1685.

Supplementary Materials

In addition to the main analyses we also tested whether colour representations that are accessed via perception and via association could be decoded using our methods. Below, we summarise these results.

Real colour analysis

For the real colour analysis, we trained the classifier to distinguish between MEG data when participants viewed the abstract shapes in different colours and tested its performance on independent real colour trials. We found that most of the colour pairs could be decoded from ~70ms after stimulus onset (Figure S1B). Yellow versus green trials could be differentiated later on in the signal (~115ms) but the colour representation was not stable (Figure 2B). Red versus orange could not be decoded, probably reflecting the high similarity between these colours (Figure S1B). Note that the decoding accuracy might be influenced by luminance differences which are smaller in the case of red versus orange and yellow versus green than in all the other colour pairs (unlike in our previous study (Teichmann et al., 2019) in which the colours were equiluminant). Peak decoding for the remaining real colour pairs was at ~135-150ms after stimulus onset (Figure S1).

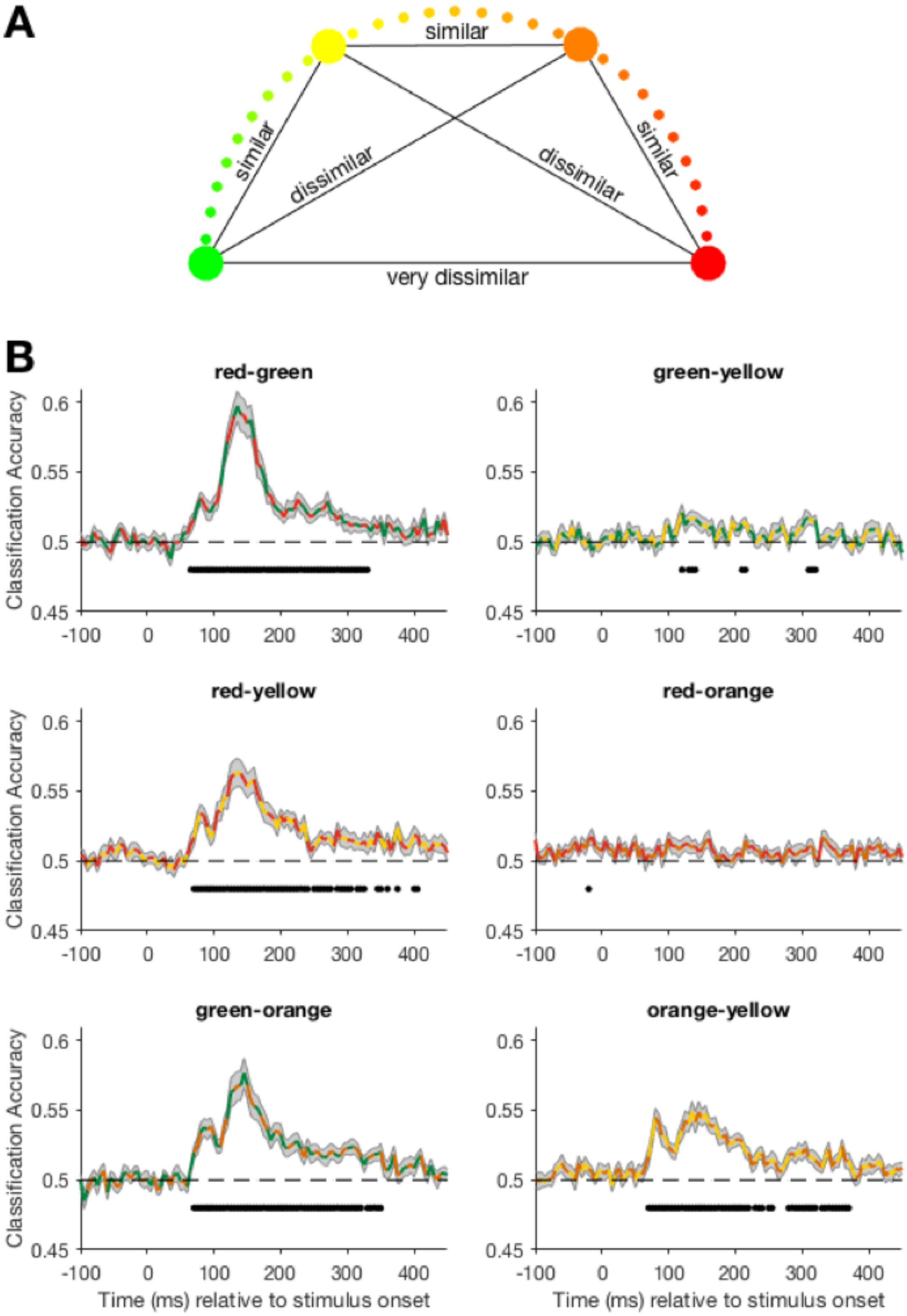

Figure S1. (A) shows the similarity between colour categories based on colour distance. (B) shows the results of the real colour decoding analysis for all colour pairs. The classifier was trained on the abstract shape trials to distinguish between two colours at a time. Each colour category contained identical shapes. The classifier was trained to distinguish between all shapes except one. Testing was done only on the shape that the classifier did not train on.

The classification accuracy over time for the real colour decoding analysis for each colour pair is plotted above. Shading represents the standard error. Black dashed line represents chance level (50% - pairwise comparison). Black dots highlight significant timepoints for each pairwise comparison ($p<0.05$), corrected for multiple comparisons.

Implied colour analysis

A controlled approach of testing whether there is any representational overlap between real and implied colours is to train a classifier on real colour and test on implied colour trials. For this analysis, there is no low-level feature such as shape or luminance that could drive the classification. Successful cross-generalisation implies that the brain representation of colour accessed via colour perception and association share characteristics. To see whether this is the case, we trained a classifier to distinguish between patterns evoked by pairs of our coloured abstract shapes, as in our first analysis. We then tested the classifier on distinguishing between the grey-scale objects that are associated with those colours. Consistent with our previous work (Teichmann et al., 2019), the representational overlap for real and implied colours dynamically evolved over time. We therefore ran this analysis as a time-generalisation analysis, training and testing the classifier at every timepoint combination (Carlson et al., 2011; King & Dehaene, 2014). We ran the analysis separately for each of the colour pairs as the real colour decoding results showed that the classifier cannot reliably distinguish all colour pairs (Figure S1). Across the time-time decoding matrices (Figure S2), we can see that the classifier can cross-generalise best between real and implied colours when the colours are most dissimilar (i.e., red and green). Accessing colour via real colour perception and implied colour activation occurred at the same time, around 150ms. For red versus green, there was additional significant decoding off the diagonal, which indicates a temporal difference in the instantiation of a similar pattern. Colour information evoked by real colours from ~150ms-450ms resembles colour information evoked by greyscale objects in a timewindow from ~150-170ms after stimulus onset. This indicates that colour information evoked by association is only briefly in the signal. There also is a reactivation of colour information for the red-green comparison at ~400ms after stimulus onset. For the red versus yellow and green versus orange contrast there is cross-generalisation in a timewindow around ~150ms after stimulus onset. In contrast, we did not observe successful cross-generalisation when training and testing on colours that are similar. This is not surprising given that we did not have a reliable model to distinguish between these real colours (i.e., red versus orange, Figure S1).

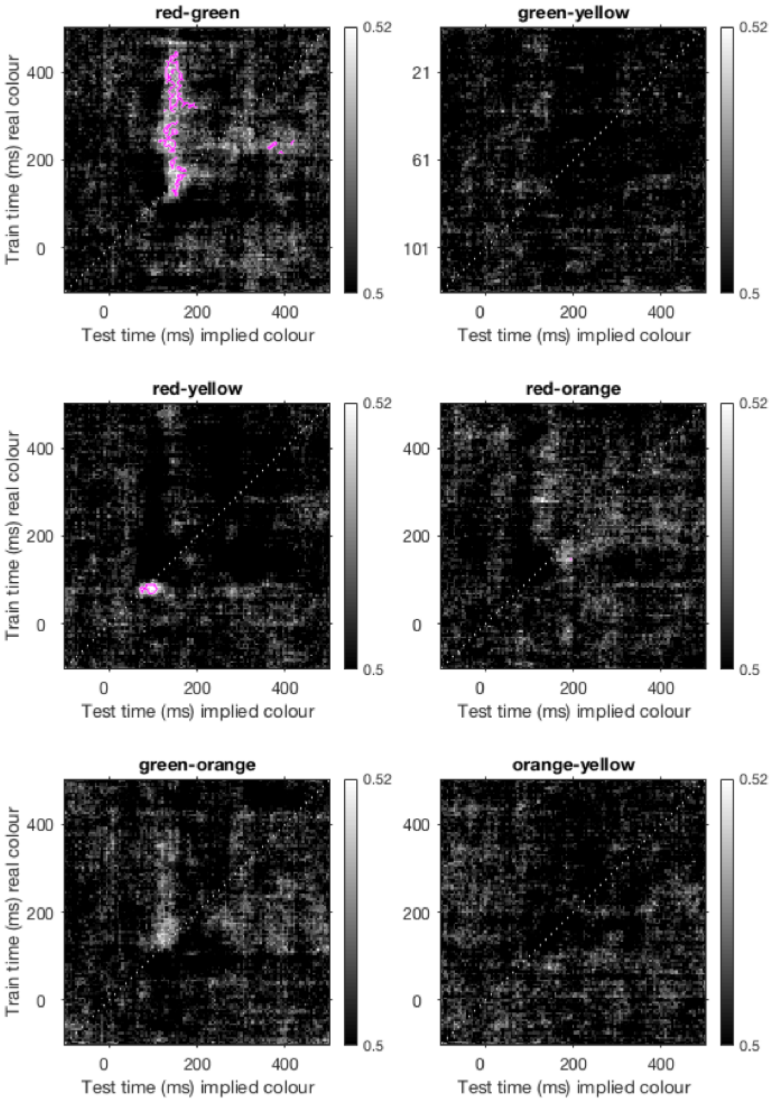

Figure S2. Time-generalisation for cross-decoding, training the classifier on real colour trials and testing on grey-scale object trials. The analysis was done separately for each pairwise colour comparison. The plots in show the decoding accuracies for every timepoint combination when training on abstract shapes and testing on greyscale object trials. Pink highlights show timepoints with decoding significantly above chance ($p<0.05$), controlled for multiple comparisons.

Overall, our results show that there are representational similarities for real and implied colours but this is only distinguishable using our methods for colours that are quite dissimilar (e.g., red and green). It is important to note that there is no colour information at all in the equiluminant greyscale object trials and that the shapes used for training the classifier are identical except in colour and luminance. That means we here have strong evidence for real and implied colour sharing an overlapping brain activation pattern that becomes apparent from around 150 to 200ms after stimulus onset, at least when the colours are dissimilar enough.