

FEATURE ARTICLE

Differential Processing of Isolated Object and Multi-item Pop-Out Displays in LIP and PFC

Ethan M. Meyers^{1,2}, Andy Liang¹, Fumi Katsuki^{3,4}
and Christos Constantinidis⁴

¹Center for Brains, Minds and Machines, Massachusetts Institute of Technology, Cambridge, MA 02139, USA, ²School of Cognitive Science, Hampshire College, Amherst, MA 01002, USA, ³Department of Psychiatry, VA Boston Healthcare System, Harvard Medical School, West Roxbury, MA 02132, USA and ⁴Department of Neurobiology and Anatomy, Wake Forest School of Medicine, Winston-Salem, NC 27157, USA

Address correspondence to email: emeyers@hampshire.edu

Abstract

Objects that are highly distinct from their surroundings appear to visually “pop-out.” This effect is present for displays in which: (1) a single cue object is shown on a blank background, and (2) a single cue object is highly distinct from surrounding objects; it is generally assumed that these 2 display types are processed in the same way. To directly examine this, we applied a decoding analysis to neural activity recorded from the lateral intraparietal (LIP) area and the dorsolateral prefrontal cortex (dlPFC). Our analyses showed that for the single-object displays, cue location information appeared earlier in LIP than in dlPFC. However, for the display with distractors, location information was substantially delayed in both brain regions, and information first appeared in dlPFC. Additionally, we see that pattern of neural activity is similar for both types of displays and across different color transformations of the stimuli, indicating that location information is being coded in the same way regardless of display type. These results lead us to hypothesize that 2 different pathways are involved processing these 2 types of pop-out displays.

Key words: attention, lateral intraparietal area, neural decoding, posterior parietal cortex, prefrontal cortex

Introduction

The ability to locate behaviorally relevant stimuli is one of the most fundamental tasks for organisms that are capable of self-directed movement, and represents a major evolutionary driving force (Karlen and Krubitzer 2007; Kaas 2012). Many species of primates have sophisticated brain systems that are used for visual search, and their abilities to locate objects visually have been extensively studied (Katsuki and Constantinidis 2014). These studies have revealed that there are 2 distinct modes of visual search (Itti and Koch 2001; Corbetta and Shulman 2002; Connor et al. 2004). One mode, called “parallel search,” occurs in displays where an object of interest is very distinct from its surroundings and it is said to “pop-out,” and leads to very fast reaction times (Treisman and Gelade 1980; Duncan and Humphreys 1989).

A second form of search, “serial search” operates on complex displays where no object stands out against its surroundings, and leads to much slower reaction times that increase with additional distracting stimuli (Wolfe and Horowitz 2004).

In parallel search, it is generally thought that pop-out elements are detected by bottom-up processes through a winner-take-all mechanism that orients attention to the most salient visual element (Koch and Ullman 1985; Itti and Koch 2001). Cortical areas representing the location of salient objects include the dorsolateral prefrontal cortex (dlPFC) and the lateral intraparietal area (LIP) (Schall and Hanes 1993; Goldberg et al. 1998; Constantinidis and Steinmetz 2001; Katsuki and Constantinidis 2012). Previous work has shown that visual firing rate responses during search propagate in a feed-forward manner from early

visual cortex to LIP and from there to dlPFC, therefore, it is thought that pop-out stimuli are first represented in parietal and then in prefrontal cortex (Buschman and Miller 2007; Siegel et al. 2015). The reverse pattern of activation is observed during serial search (Buschman and Miller 2007; Cohen et al. 2009).

There are, however, a few shortcomings of a simple dorsal stream feed-forward account of parallel search. In particular, behavioral results have shown that reaction times in parallel searches can be affected by the recent history of the search task and by stimulus expectations (Wolfe et al. 2003; Leonard and Egeth 2008), which indicates that a simple bottom-up account of pop-out attention is not adequate (Awh et al. 2012). Similarly, recent computational modeling work has suggested that the neural processing underlying parallel search might involve more than feed-forward processing (Khorsand et al. 2015).

To examine the neural basis of parallel search, we analyzed data from an experiment in which monkeys needed to remember the location of a salient object in trials that consisted of 2 types of pop-out displays; the first type of displays, which we call “isolated cue displays,” consist of single object, while the second type of displays, which we call “multi-item displays” consist of an object of one color in the midst of several distractor objects of a different color (Katsuki and Constantinidis 2012). By precisely characterizing the time course of firing rate increases and the time course of spatial information increases we were able to detect significant differences between how LIP and dlPFC process these 2 types of displays. Based on these results we hypothesize that there are 2 pathways involved in parallel search which helps to account for several other findings in the literature.

Materials and Methods

All animal procedures in this study followed guidelines by the U.S. Public Health Service Policy on Humane Care and Use of Laboratory Animals and the National Research Council’s Guide for the Care and Use of Laboratory Animals, as reviewed and approved by the Wake Forest University Institutional Animal Care and Use Committee.

Experimental Design and Neural Recordings

Two male rhesus monkeys (*Macaca mulatta*) engaged in a delayed match-to-sample task where the monkeys needed to remember the spatial location of a cue stimulus (Fig. 1). Trials began when the monkeys fixated on a 0.2° white square in the center of the monitor and pulled a behavioral lever. After a trial was initiated, the cue stimulus was presented at 1 of 9 possible locations on a 3 × 3 square grid. The cue was 1.5° in size and could be either red or green. The 3 × 3 grid was set up such that adjacent locations on the grid were separated by 15° of visual angle. On isolated cue trials, the cue was shown on a blank background. On multi-item display trials, the cue was shown surrounded by 8 distractors of the opposite color that filled out all of the other locations in the 3 × 3 grid. In the majority of experimental sessions, the cue appeared in 1 of only 4 of these locations. The location and color of the cue were randomized from trial-to-trial. Isolated cue and multi-item display trials were also randomly interleaved. The luminance of cue and distractor stimuli of either color was 24 cd/m²; the luminance of the background was <0.1 cd/m².

After being presented for 500 ms, the cue was removed and a delay period that consisted of a blank screen was presented for 1000 ms. A sequence of 0–2 nonmatch stimuli was then

presented with each stimulus lasting for 500 ms followed by a 500 ms blank screen delay. The monkeys received a juice reward for releasing the lever within 500 ms of a stimulus that matched the same location as the cue. Releasing the lever at any other point in the trial, failing to release the lever within the 500 ms window, or making an eye movement outside the central 2° of the monitor, were considered errors and resulted in an aborted trial.

Single-unit recordings were made from the monkeys from the dlPFC (areas 8a and 46) and from area LIP in the lateral bank of the intraparietal sulcus while the monkeys engaged in the task. A total of 408 neurons were recorded in LIP (316 from M1, and 92 from M2) and 799 neurons were recorded from PFC (643 from M1, and 156 from M2); between 1 and 20 neurons were recorded simultaneously (median = 6 neurons). Recordings were performed from the right hemisphere, in both areas and for both monkeys. In monkey M1, 148 PFC neurons were recorded initially. In monkey M2, recordings from 14 LIP neurons were performed initially. Recordings from LIP and PFC were interleaved from that point on, including in sessions with recordings from both areas simultaneously. The data used in these analyses were previously presented (Katsuki and Constantinidis 2012) and more details about the experimental design and neural recordings can be found there. A reaction time version of this task was also conducted in the original study by Katsuki and Constantinidis (2012) which demonstrated that these types of displays show the typical behavioral pattern for parallel search in monkeys, where reaction times are not influenced by the number of distractors present.

Data Analysis: Decoding Analysis

A population decoding approach was used to analyze the data in this article (Meyers et al. 2008, 2012). Briefly, a maximum correlation coefficient classifier was trained on firing rates of a population neurons to discriminate between the location (Figs 2–6) or the color (Figs 7 and 8) of the cue stimulus. The classification accuracy was then calculated as the percentage of predictions that were correctly made on trials from a separate test set of data. All analyses were done in MATLAB using versions 1.2 and 1.4 of the Neural Decoding Toolbox (Meyers 2013), and the code for the toolbox can be downloaded from www.readout.info. In this article, we use the term “decoding accuracy” and “information” interchangeably since decoding accuracy gives a lower bound on the amount of mutual information in a system (Quiñ Quiroga and Panzeri 2009) (both terms refer to the ability of neural activity to differentiate between different experimental conditions, in contrast to firing rate increases, which could increase same way for all experimental conditions). For more in depth discussion of the decoding analyses see Meyers and Kreiman (2012) and Quiñ Quiroga and Panzeri (2009).

Only data from trials in which the monkey performed the task correctly were used in these analyses. When decoding the location of the cue (Figs 2–6) we constrained our analyses to decoding only 4 locations of the cue (upper center, middle left, middle right, and lower center) since many experimental sessions only had the cue presented at these locations. All neurons that had recordings from at least 5 correct trial repetitions with the cue appearing at these 4 locations for the red and green isolated and multi-item trials were included in all our decoding analyses. This led to a population of 651 PFC and 393 LIP neurons. From these larger populations, pseudopopulations of 350 randomly selected neurons were created. Trials from

these pseudopopulations were randomly split into training and test sets using 5-fold cross-validation, where the classifier was trained on data from 4 splits and tested on the fifth split, and the procedure was repeated 5 times using a different test split each time. For each cross-validation split, the mean and standard deviation of the firing rates of each neuron were calculated using data from the training set, and all neuron's firing rates were z-score normalized based on these means and standard deviations, before the data was passed to the classifier. This procedure was repeated over 50 resample runs, where different random pseudopopulations and training and test splits were created on each run. The final results reported are the average classification accuracy over these 50 runs and over the 5 cross-validations splits (because the temporal-cross decoding plots in Figure 5 were computationally expensive to create, only 10 resample runs were used in these figures).

When decoding location of the cue (Figs 2–5) each cross-validation (CV) split contained one example from red and green cue trials at each of the 4 locations (8 points per CV split: $8 \times 4 = 32$ training points and 8 test points). Because we were interested in comparing the latency of location information in LIP and PFC (and because the decoding accuracy was high), we used 30 ms bins sampled at 1 ms resolution for the cue location decoding analyses in order to get a more precise temporal estimate for when information was in a brain region. When assessing the dynamics of the population code (Fig. 5), we applied a temporal-cross decoding analysis (also called the temporal generalization method) where we trained the classifier on 4 trials from each condition at one time bin and then tested the classifier with data from a different trial from either the same time bin or from a different time bin (Meyers et al. 2008; King and Dehaene 2014).

When decoding whether the spatial information was represented in a color invariant manner (Fig. 6), we applied a generalization analysis (Meyers 2013) where we trained a classifier on 4 trials from 1 of our 4 conditions (red/green cue, single/multi-item displays) and then tested the classifier with either a fifth trial of the same condition or a from a trial from a different condition (this led to 4 points per CV split: 4 locations \times 4 repetitions = 16 training points, and 4 test points). When decoding color information (Figs 7 and 8), each cross-validation split contained a red and green trial from each of the 4 locations (8 points per CV split: 32 training points and 8 test points). For the color invariant and color decoding analyses (Figs 6 and 7) we used a larger 100 ms bins because decoding accuracy is higher and less noisy with larger bin sizes, which makes it easier to detect when information is present (albeit at the cost of temporal precision). This was useful for the color invariant decoding because there was less training data in this analysis, and was particularly useful for the color decoding analyses since the decoding accuracies were overall very low/noisy when decoding color information, although for the color decoding in Figure 8, we again used a 30 ms bin to more precisely compare the timing of location and color information (in the past we have used bin sizes in the range of 100–500 ms; Meyers et al. 2008, 2012, 2015; Zhang et al. 2011). For all analyses, the decoding results were always plotted relative to the time at the center of the bin.

When comparing the color decoding accuracies to the spatial information decoding accuracies (Fig. 8b), we used a normalized rank decoding measure (Mitchell et al. 2004) in order to put the results on the same scale since chance for decoding spatial information was 25% while chance for decoding color was 50%. Normalized rank results calculate where the correct prediction is on an ordered list of predictions for all stimuli—where 100%

indicates perfect decoding accuracy, 50% is chance, and 0% means one always predicted the correct stimulus as being the least likely stimulus.

To assess when the decoding accuracy was above what is expected by chance, we used a permutation test. This test was run by randomly shuffling the labels of which stimulus was shown in each trial separately for each neuron, and then running the full decoding procedure on these shuffled labels to generate one point in a “shuffled decoding” null distribution. This procedure was then repeated 1000 times to generate a full shuffled decoding null distribution for each time bin. The solid bars on the bottom of Figures 2, 3, 7 and 8 indicate time points in which the decoding results were greater than all 1000 points in the shuffled decoding null distribution ($P < 0.001$ for each time bin). The thin traces in Figure 2 show the maximum and minimum values from the shuffled decoding null distribution for each time bin. More information about this permutation test can be found in (Meyers and Kreiman 2012). When reporting the “first detectable information latencies” (Table 1), a 30 ms bin was used and the latencies reported are relative to the end of these bins (e.g., if the firing rates were computed over a bin ranging from 71 to 100 ms relative to the stimulus onset, then a latency of 100 ms for this bin would be reported in the article). These latencies were defined as the first time of 5 consecutive bins in which the real decoding accuracy result was greater than all points in the shuffled decoding null distribution. The choice of using 5 consecutive bins was chosen prior to doing the analysis, however, from looking at the results we see that once the results became statistically significant they remained highly statistically significant and so this choice did not affect the results.

To assess whether there was a difference in the latency when location information first appeared in LIP compared with dlPFC we ran another permutation test. This test was done by combining the data from the LIP and dlPFC populations of neurons into a single larger population, and then randomly selecting 651 neurons to be a surrogate dlPFC population and taking the remaining 393 neurons to be a surrogate LIP population; that is, we created surrogate populations that were consistent with the null hypothesis that there is no information latency difference between LIP and dlPFC so neurons in the LIP and dlPFC populations are interchangeable. We then ran the full decoding procedure on these surrogate LIP and dlPFC population pairs and we repeated this procedure 1000 times to get 1000 pairs of decoding results that were consistent with the null hypothesis that LIP and dlPFC had the same information latency (Fig. S3 shows the decoding accuracies for the LIP and PFC results and all the surrogate runs). For each of these 1000 pairs of surrogate LIP and PFC decoding accuracies, we calculated their first detectable information latency, and we subtracted these latencies to get a null distribution of latency differences. A P -value was then created based on the number of points in this null distribution of latency difference that were as large or larger than the real LIP – PFC latency difference (this is the P -value reported in the body of article).

Additionally, to make sure that the results were robust, we assessed the latency difference at a range of decoding accuracy threshold levels (rather than just using the decoding accuracy level at the first detectable latency difference). To do this we calculated the time when the decoding first reached a 45% accuracy level, we will refer to these latencies as $L45_{PFC}$ and $L45_{LIP}$ (a 45% decoding accuracy that was greater than all the points in the shuffled decoding null distribution for all time bins, and thus corresponded to a $P < 0.001$ in our shuffled decoding permutation test). We then calculated $|L45_{PFC} - L45_{LIP}|$

to get the real latency difference between LIP and PFC at this 45% accuracy level, and we also calculated the $|L45_{PFC} - L45_{LIP}|$ for all 1000 surrogate pairs to get a null distribution of latency differences under the assumption that LIP and PFC had the same latency. A *P*-value was then calculated as the proportion of points in this null distribution that was greater than the actual $|L45_{PFC} - L45_{LIP}|$ difference. To make sure this procedure was robust, we repeated it for different decoding accuracy levels from 45% to 65% (i.e., we assessed $|L45_{PFC} - L45_{LIP}|$, $|L46_{PFC} - L46_{LIP}|$, ..., $|L65_{PFC} - L65_{LIP}|$) calculating the *P*-value at each decoding accuracy level and we plotted these *P*-values as a function of the decoding accuracy level in Figure S3e, S3i.

Data Analysis: Firing Rate Analyses

To calculate population firing rates (Figs 4, S1, and S6) we simply averaged the firing rates of all neurons from the relevant trials together. Only neurons that went into our decoding analysis were used when calculating these mean firing rates in order to make a direct comparison to the decoding results (see selection criteria above). Because we wanted to report results in meaningful units of spikes per second, we did not normalize the firing rates of neurons before averaging. However, we checked all results to make sure that neurons with higher firing rates were not unduly influencing the results by applying a *z*-score normalization to each neuron before averaging all the neurons together, and saw that the results looked very similar and led to all the same conclusions. When comparing firing rate to decoding results we scaled the axes so that the maximum and minimum heights of firing rates and decoding results are the same, which makes it much easier to compare their time courses of these different results. Similarly, when comparing firing rates from LIP to dlPFC we also scale the axes to have the same maximum and minimum heights when plotted to make a comparison of their time courses possible (as can be seen on the figures axes, LIP has a higher population firing rate than dlPFC).

To compare firing rates when the cue stimulus was shown inside a neuron's receptive field versus outside of a neuron's receptive field for the multi-item display trials (Fig. S6a, S6b), we recreated the Figure 3a, b of Katsuki and Constantinidis (2012) (for our analysis we only used neurons that were used in the other analyses in our study; i.e., that had at least 5 repeats of all the stimulus conditions which was as slightly different set of neurons than used by Katsuki and Constantinidis (2012)). To do this analysis we found all neurons that showed spatial selectivity as determined by an ANOVA with an alpha level of 0.01 using the average firing rate in a 500 ms bin of data from the "isolated cue trials." For each of these neurons, we determined the spatial location that had the highest firing rate as the neuron's "preferred location" (i.e., the location in the neuron's receptive field), and we chose the location that has the lowest firing rate as the neuron's "anti-preferred location" (i.e., the location that was outside the neuron's receptive field) again using data from the isolated cue trials. We then found the firing rates on the "multi-item display trials" when the cue was at the "preferred" location and averaged these firing rates from all neurons together and did the same procedure for the "anti-preferred" location. The analysis was done separately for LIP (Fig. S6a) and for dlPFC (Fig. S6b) and we combined the data from the red cue and green cue trials. An analogous procedure was used for plotting the firing rates inside and outside of the receptive field for the isolated object trials (Fig. S6c, S6d), but for this analysis we found each neuron's preferred and anti-preferred locations using data from the multi-item display trials, and then we plotted the population average firing rate on the isolated object trials based on those locations.

Results

In order to gain insight into the neural processing that underlie parallel search, we analyzed neural activity in LIP and dlPFC that was recorded as monkeys performed a delayed match to sample task (Fig. 1). Monkeys were shown either

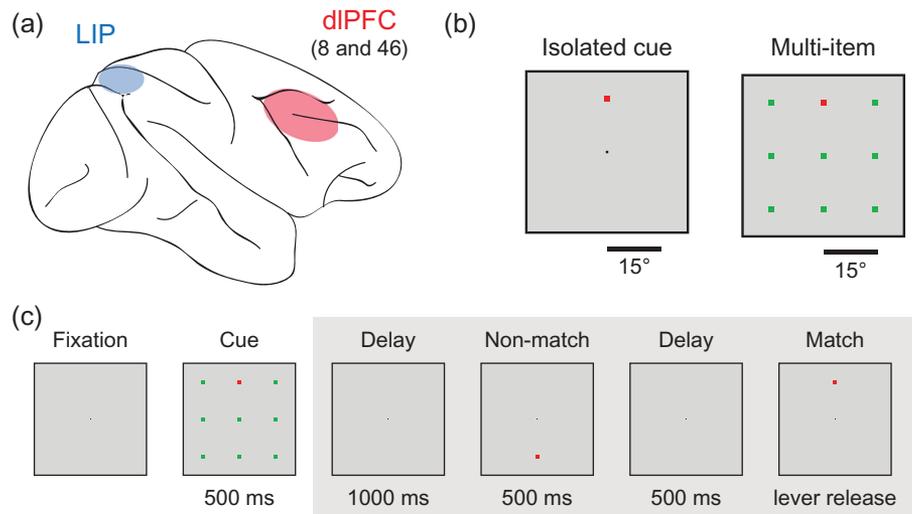


Figure 1. Brain regions and experimental design. (a) Locations of the 2 brain regions (LIP and dlPFC) where the recordings were made. (b) Examples of the isolated cue and multi-item displays. The stimuli were 1.5° in size and were displayed 15° apart. The cue was either green or red and the distractors were the opposite color and the cue and the distractors were isoluminant. (c) An example of an experimental trial. Trials were initiated when a monkey fixated on a dot in the center of the screen. A cue, which consisted of either a single square displays on a blank background (isolated cue trial; not shown) or a square of one color with 8 other "distractor" squares of a different color (multi-item trial) was displayed for 500 ms. The monkey needed to remember the location of the cue and perform a delayed match to sample task on the cue location. All analyses in this article only examine data from the time when the cue display was displayed (i.e., data from time periods in the gray box were not analyzed).

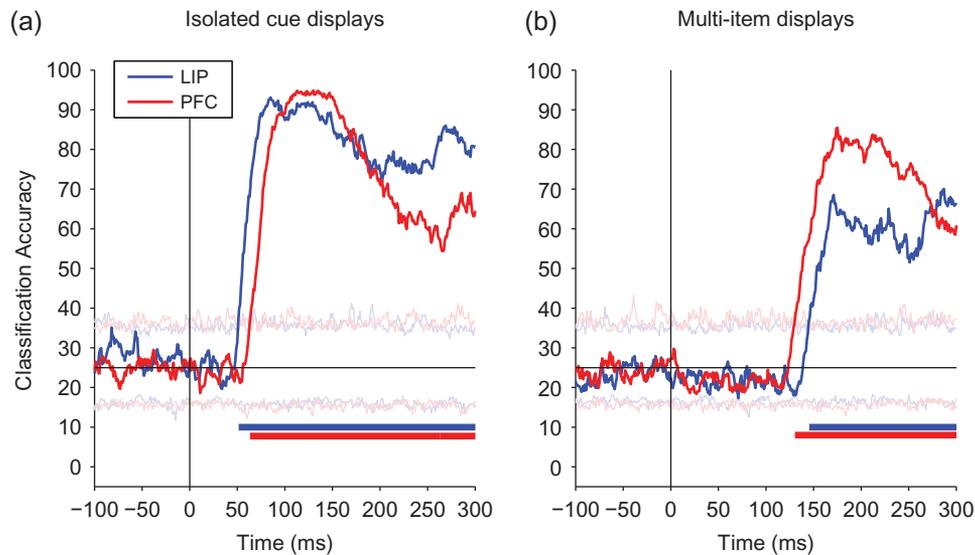


Figure 2. Comparison of decoded information about the cue location from LIP and PFC. A comparison of information decoded from LIP (blue) and PFC (red). (a) Isolated cue trials, showing information increases earlier in LIP than in dIPFC. (b) Multi-item trials, showing information increases earlier in PFC than in LIP. The black vertical line shows the time of stimulus onset, and the black horizontal line shows the level of decoding accuracy expected by chance. Solid bars at the bottom of the plots show when the decoding accuracies for LIP (blue) and PFC (red) are above chance, and light horizontal traces show the maximum and minimum decoding accuracies obtained from the null distribution.

isolated cue displays where a single stimulus appeared by itself, or multi-item displays where a salient cue stimulus was shown surrounded by other stimuli of opposite color (Fig. 1b). The monkeys needed to remember the location of the salient stimuli and release a lever when a second stimulus was shown at the same location (Fig. 1c). The cue stimulus could be either red or green, and on the multi-item display trials the distractor stimuli were of the opposite color as the cue. The monkeys were able to process these displays in parallel rather than requiring serial search; when tested with a version of the task requiring an immediate response, using identical stimuli, reaction time was flat as a function of number of elements in the display (Katsuki and Constantinidis 2012). Our analyses here focused on the first 500 ms of activity after the cue was shown since this was the time window in which information first entered these brain regions.

Location Information Appears Much Later in Multi-item Display Trials Relative to Isolated Cue Trials

Previous work has shown that firing rates are propagated in a feed-forward manner from LIP to dIPFC (Buschman and Miller 2007; Katsuki and Constantinidis 2012; Siegel et al. 2015), and indeed, a firing rate analysis of our data shows that firing rates increases also first occur in LIP and then in PFC for both the isolated and multi-item display trials (Fig. S1). However, while it has often been assumed that information increases have the same time course as firing rate increases (Nowak and Bullier 1997), this assumption does not necessarily need to be true (Heller et al. 1995), and since we are generally interested in how the brain processes information to solve tasks, assessing the latency of information should be most relevant for our understanding. Thus, we examined the time course of information about the location of the cue stimulus using a decoding analysis, where a classifier had to predict the spatial location of the cue stimulus from the firing rates of populations of neurons in LIP and dIPFC (see Materials and Methods). To do this analysis we combined data from the red and green cue trials and tried

Table 1 Summary of first detectable information latencies for LIP and dIPFC for the isolated cue and multi-item displays. The first detectable information latency was defined as the first of 5 bins when the real decoding result was larger than all the values in the null distribution. The results are based on using a 30 ms bin, and the time reported is the end of the bin (see Materials and Methods for more details). For the isolated cue trials, information first appeared in LIP prior to dIPFC. For the multi-item display trials, the latency in both areas was delayed and information now first appeared in dIPFC prior to LIP

	Isolated cue	Multi-item display	Difference: multi-item – isolated
LIP	67	161	94
dIPFC	79	146	67
Difference: LIP – dIPFC	-12	15	

to predict which of 4 locations the cue stimulus was presented at (Supplemental Fig. 2 shows the results separately for the red and for the green cue trials, and we analyze the color information later in the article).

Results from these analyses on the isolated cue trials (Fig. 2a) showed that information appeared ~12 ms earlier in LIP than in dIPFC, and a permutation test revealed that these results were statistically significant ($P < 0.01$, also see Fig. S3a–e). This pattern was consistent for both the red and green cue trials (Fig. S2a, S2b). Thus these results suggest that information (as well as firing rates) propagate in feed-forward manner from LIP to dIPFC (Table 1). Surprisingly, however, for the multi-item trials information appeared ~15 ms later in LIP compared with dIPFC (Fig. 2b), and a permutation test revealed that these results were statistically significant ($P < 0.01$, also see Fig. S3f–j). This pattern of results were again consistent across the 2 independent data sets for the red and green cue trials and were robust to analysis parameters such as bin size (Fig. S2c, S2d). Thus, these results are consistent with information flowing from dIPFC to LIP for the multi-item displays (or potentially both

regions getting the information from a common brain region with a larger delay for LIP). Additionally, the small latency differences between LIP and dlPFC suggest that the information might be traveling directly between these regions, which is consistent with the known anatomy (Cavada and Goldman-Rakic 1989; Felleman and Van Essen 1991).

In Figure 3, we also replot the results of Figure 2 to show a direct comparison of the time course of information on the isolated cue trials to the multi-item display trials. In contrast to the small latency difference seen when comparing LIP to dlPFC, information was substantially delayed in the multi-item display trials relative to the isolated cue trials on the order of 94 ms for LIP and 67 ms for dlPFC (Table 1). Thus these findings suggest that location information on the isolated cue trials is traveling in a “bottom-up” direction from LIP to dlPFC while location information in the multi-item trials takes much longer, perhaps due to processing requirements, and appears to be traveling in the reverse direction from dlPFC to LIP. We explore possible explanations for this substantial delay of spatial information later in the article.

To better compare the time course of firing rates with the time course of decoded information, we also plotted the firing rates and decoded spatial information on the same figure (Fig. 4). These results show that in the isolated cue trials for both LIP and dlPFC (Fig. 4a, b), the time course of firing rate increases and increases in decoding accuracy are closely aligned indicating that spatial information appears as soon as there is an increase in firing rate. In contrast, for both LIP and dlPFC in the multi-item trials, the population firing rate increases much earlier than the increase in spatial information (Fig. 4c, d). Thus the firing rate increases are much more closely tied to the onset of the stimulus rather than to the increase in information about the location of the cue.

Examining Dynamic and Stationary Neural Coding

The fact spatial information is delayed relative to the time course of firing rate increases in the multi-item displays, but not in the isolated cue displays, raises questions about the role

of the early firing rate increase in the isolated cue displays. One possible explanation is that the initial spatial information seen in the isolated cue displays is tied to lower level features (e.g., “there is retinotopic stimulation at location X”) while later activity might be coding an abstract representation of spatial information (e.g., “location X contains an object of interest, even though there is visual stimulation throughout the visual field”). If this is the case, then decoding spatial information would work from the onset in the isolated cue trials because the classifier could use activity related to simple lower level retinotopic features, but decoding would not work for the multi-item trials because these lower level features would activate neural activity at all locations. As a first test of this idea, we applied a “temporal generalization analysis” where a classifier was trained with data from one point in time in the trial, and then tested on data from a different point in time (Meyers et al. 2008, 2012; Stokes et al. 2013; King and Dehaene 2014). If the information contained in the initial response in the isolated cue trials was different then the information contained later, then training the classifier on the initial firing rate increase in the multi-item trials would lead to poor classification performance later in the trials (King and Dehaene 2014).

Results from LIP reveal a square region of high decoding accuracies in the temporal-cross decoding plot (Fig. 5a), indicating that spatial information is represented the same way throughout the first 500 ms of the trial, although the initial representation in both areas showed a slightly more reliable than later in time, which can be seen by the slightly higher decoding accuracies when information first appears. Similar results are seen in the isolated decoding results in dlPFC (Fig. 5b). Applying this temporal generalization analysis to results to the multi-item trials showed a similar stationary code as seen in the isolated cue trials (Fig. 5c, d). Examining the results over a longer time window also revealed that this stationary pattern of results extended across the entire delay period (Fig. S4), which is a similar stationary code to what has been reported in a few other studies (Zhang et al. 2011; Murray et al. 2017). Overall, given these results show that spatial information is being coded similarly at all time periods, although

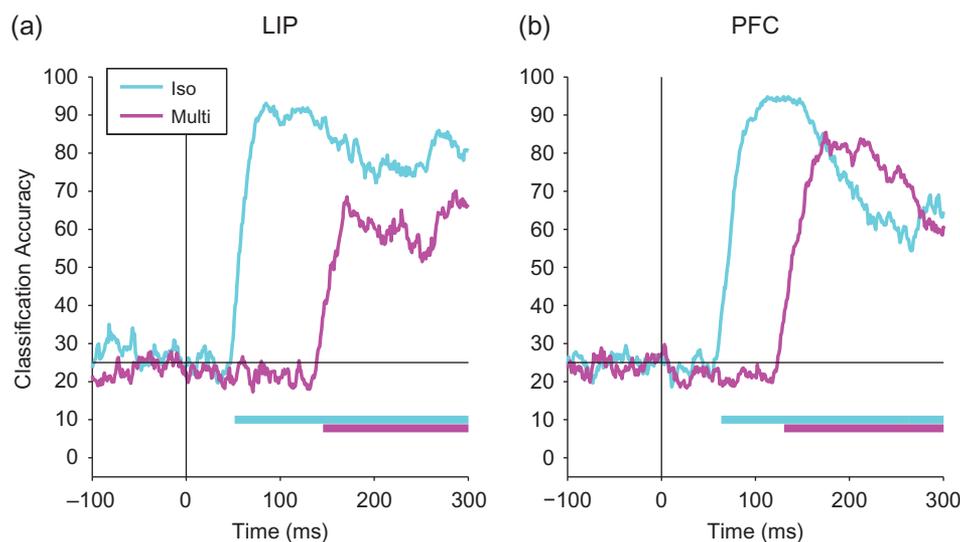


Figure 3. Comparing decoding accuracies for the isolated cue and multi-item trials. A comparison decoding accuracies for isolated cue (cyan) and multi-item trials (magenta) for LIP (a) and PFC (b). Information appears much early on the isolated cue trials compared with the multi-item display trials. The black vertical line shows the time of stimulus onset, and the black horizontal line shows the level of decoding accuracy expected by chance. Solid bars at the bottom of the plots show when the decoding accuracies for isolated cue trials (cyan) and multi-item trials (magenta) are above chance.

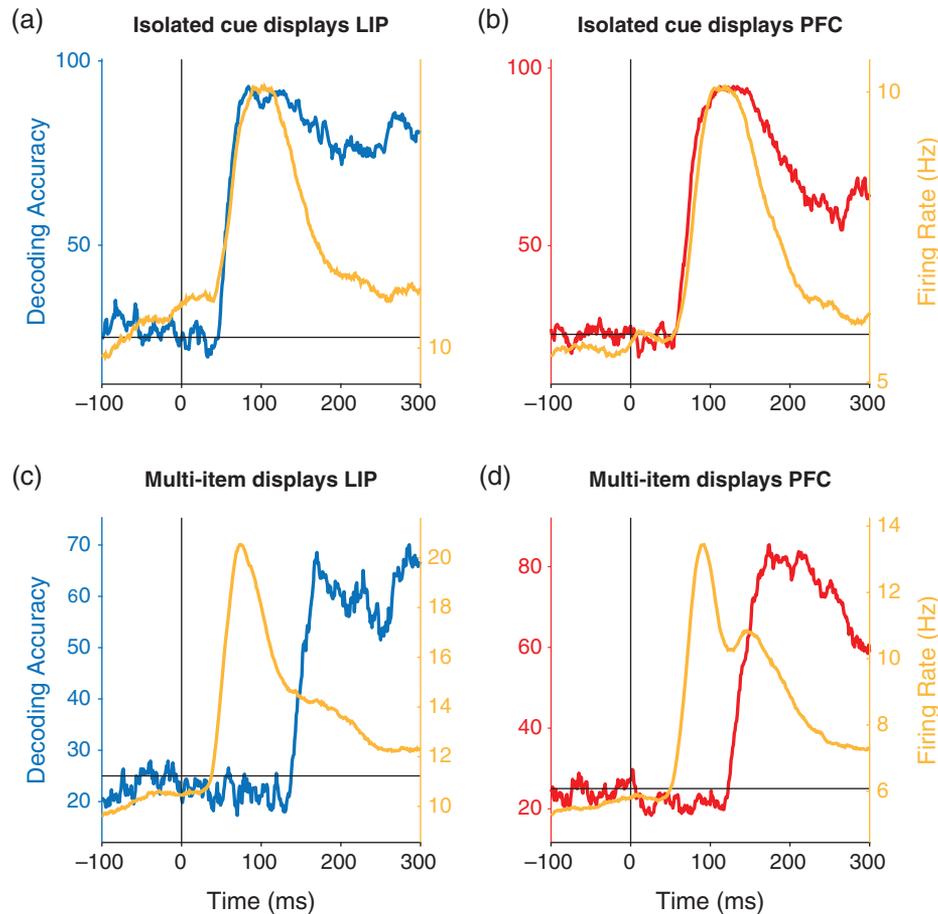


Figure 4. Comparison of decoding accuracy and firing rates. The yellow traces show the average population firing rates and the blue and red traces show the decoding accuracies for LIP (left/blue) and PFC (right/red). (a, b) Results for the isolated cue trials. The decoding information and population firing rates have the similar time courses. (c, d) Results for the multi-item displays trials. Firing rate increases occurs much earlier than the decoded information increases.

the higher decoding accuracy in the initial response could be related to information about lower level features.

Spatial Location Information is Invariant to the Color of the Cue

To further examine whether the initial information in LIP and dlPFC represented abstract spatial locations as opposed to low-level visual features, we examined whether the representation of spatial information was invariant to the color of the cue stimulus, and whether spatial information was represented the same way in isolated cue trials compared with multi-item display trials. To test this idea we conducted a “generalization analysis” (Meyers 2013) where we trained classifiers using data from trials from each of the 4 conditions: (1) isolated red cue displays, (2) isolated green cue displays, (3) multi-item red cue displays, and (4) multi-item green cue color displays. We then tested these classifiers using data from isolated red cue trials (using a different subset of data than was used to train the red isolated cue classifier). If the representations are invariant to the color of the cue stimulus, then training and testing on data from different cue colors should achieve a decoding accuracy that is just as high as when training and testing a classifier with data from the same cue color condition (e.g., training on isolated green cue trials and testing on isolated red cue trials,

should yield the same performance training and testing on the isolated red cue trials).

Results from LIP are shown in Figure 6a. The red trace shows the results from training a classifier with the isolated red cue trials and also testing with data from the isolated red cue trials which serves as a control case (this is same as the blue trace in Fig. S2a). The blue trace in the figure shows the color invariant results for isolated cue displays, that is, training with data from the isolated green cue trials and testing again with data from the isolated red trials. As can be seen, the results generalize almost perfectly across color of the cue for the isolated cue displays. The cyan trace in Figure 6a shows the results evaluating whether the neural representing is invariant to distractors, that is, training the classifier on the multi-item display red cue trials (with green distractors), and again testing the classifier with the isolated red cue trials. The information in this case is delayed, as discussed above, however, overall the classification accuracy is robust to the presence of the green distractors. Finally, the green trace shows the results from training with data from the multi-item green displays trials (and red distractors) and again testing with data from the isolated red cue trials. The results are very similar to the results when training with the multi-item display red cue trials—namely the information is delayed as expected, and the decoding accuracy levels are nearly identical. This result is particularly impressive because the trained spatial location was based on a green cue

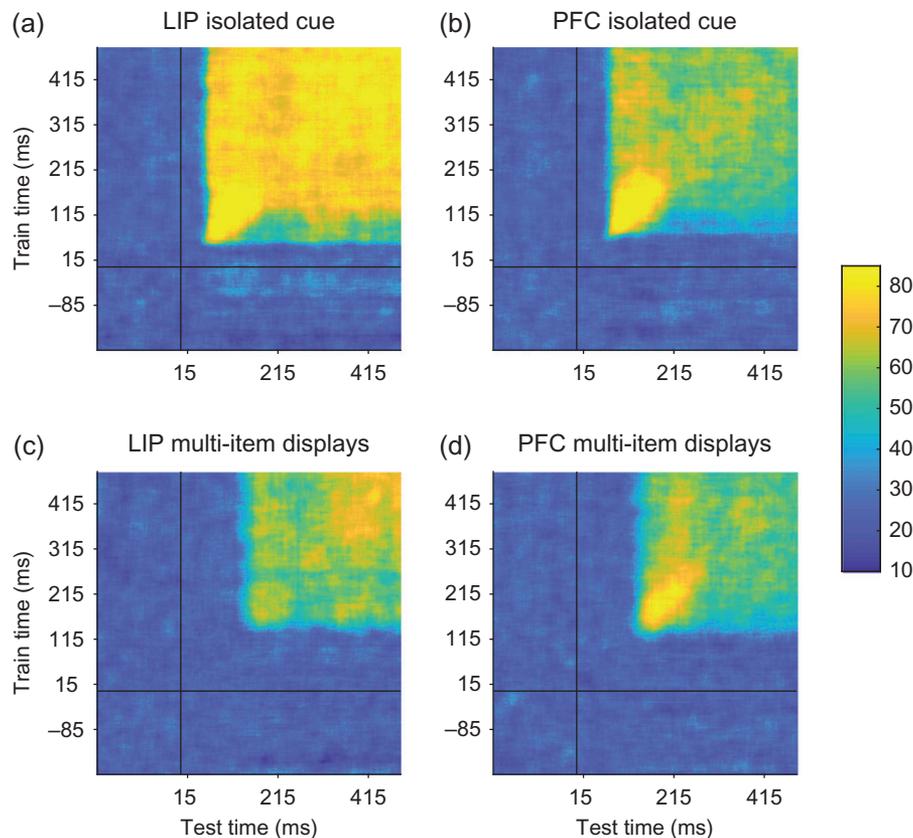


Figure 5. Temporal-cross decoding plots examining dynamic population coding. The temporal-cross decoding plots for LIP and PFC isolated cue trials (a and b, respectively), and LIP and PFC for the multi-item display trials (c and d, respectively). Overall training and testing the classifier at different points in time led to similar levels of performance as when training and testing the classifier at the same point in time, although there was often a slightly higher level of decoding accuracy when information first entered these brain regions. Thus information was contained in a relatively static code, at least compared with previous studies (Meyers et al. 2008, 2012; Stokes et al. 2013).

that is the opposite color as the red test cue color, while simultaneously there were red distractors at irrelevant spatial locations that match the red color of the test cue that needed to be decoded. Thus, this is strong evidence that LIP contains an abstract representation of location information that is invariant to the color of the cue.

The same pattern of results was seen in dlPFC (Fig. 6b) showing that dlPFC also contains an abstract representation of spatial information. An analysis based on testing the classifier with multi-item red cue trials (with green distractors) shows the same pattern of results for both LIP and PFC (Fig. 6c, d), which make an even stronger case that the data is invariant to both the color of the cue and across distractors. Finally, the same pattern of results is seen when testing the classifier using green cue trials (Fig. S5).

Color Information is Present in LIP and dlPFC on the Multi-item Trials Although it is Relatively Weak

The previous results show that the spatial information in LIP and dlPFC is invariant to the color of the cue. However, it is still possible that color information is present in the population of LIP or dlPFC neurons. Prior work has shown it is possible to decode information in a way that is invariant along a particular dimension, while also being able to decode information about that same dimension; for example, from the same population of neurons it is possible to decode information about the

identity of an individual face regardless of the pose of the head, and also to decode the pose of the head regardless of the individual face (Hung et al. 2005; Meyers et al. 2015). Color information is critical for resolving the location of the cue in the multi-item display trials, so if LIP or dlPFC is performing the computation that converts low-level visual features into an abstract spatial representations, then color information should be present in these populations prior to the presence of spatial information.

To analyze whether color information was present in LIP and dlPFC, we trained a classifier to discriminate trials when a red cue was shown from trials when a green cue was shown using data in 100 ms bins. Figure 7a shows that the ability to decode color in isolated cue trials was consistently at chance for both LIP and dlPFC indicating that there was little color information in either area for the isolated cue trials. In contrast, results from the multi-item display trials in LIP (Fig. 7b) revealed that there was a small amount of color information present in LIP and dlPFC that was above chance ($P < 0.001$ permutation test). Figure 8 compares the time course of color information to the time course of location information on the multi-item trials using smaller 30 ms bins. While overall the decoding accuracy for color information is much lower and noisier than the decoding accuracy for cue location information, the results do show that first detectable color information had a latency of 108 ms for LIP and 143 ms for PFC which is earlier than the reliably detectable location information of 161 in

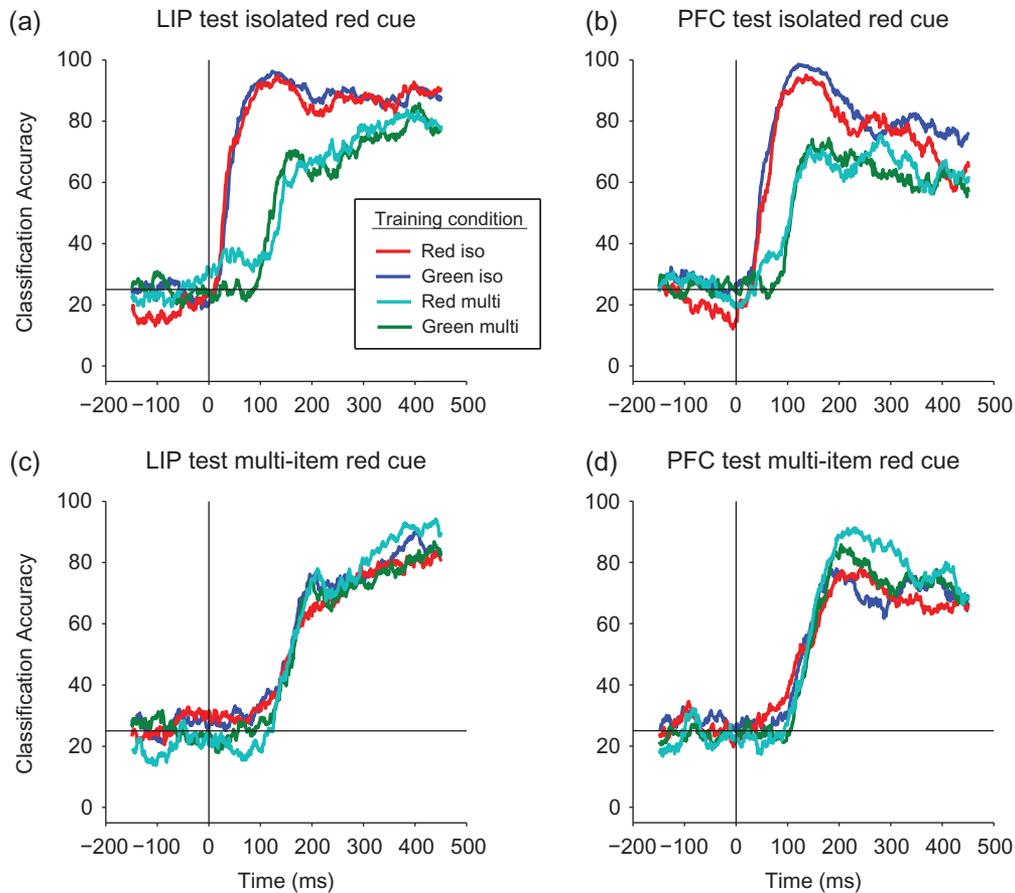


Figure 6. Examining whether the spatial location decoding generalizes across the color of the cue. (a, b) Results from LIP (a) and from PFC (b) when testing with isolated red cue data. The color lines in the figure legend show the different conditions under which the classifier was trained: (1) red traces: the classifier was trained on (different) isolated red cue trials; (2) blue traces: the classifier was trained on isolated green cue trials. The performance of the classifier trained on red cue trials was almost identical to the performance of a classifier trained on green cue trials showing spatial information is highly color invariant for the isolated cue trials. We also trained the classifier on multi-item trials: (3) cyan traces: training the classifier with multi-item red cue trials; and (4) green traces: training the classifier with multi-item green cue trials (again all testing was done on the isolated red trials). Since information on the multi-item trials appears later, the results are shifted in time, however, the classification accuracy was still highly insensitive to color, even when training for the location using a green cue among the red distractors. (c, d) Results from LIP (c) and PFC (d) when testing on data from red cue multi-item trials and training on the all 4 cue/distractor combinations. Again the results are highly invariant to the color of the cue.

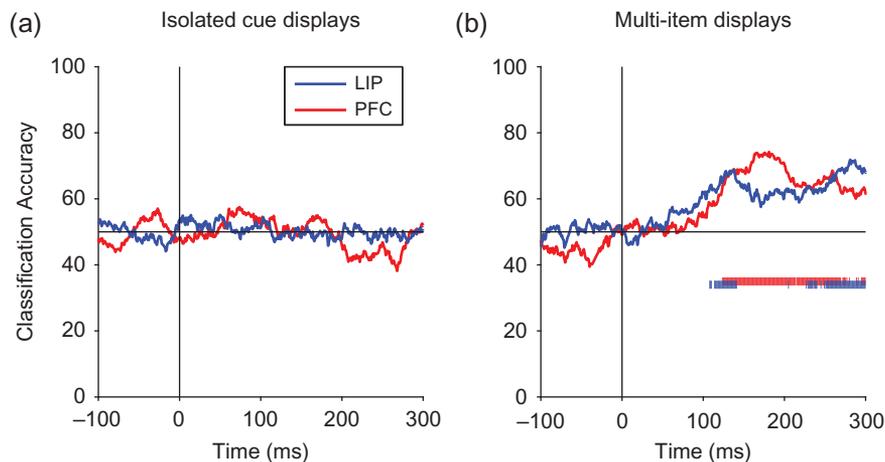


Figure 7. Decoding color information for LIP and PFC. Results from decoding the color of the cue stimuli (red vs. green) for the isolated cue and multi-item display trials. Solid bars at the bottom of the plots show when the decoding accuracies for LIP (blue) and PFC (red) are above chance. (a) For the isolated cue displays, the decoding accuracy for the color information was never significantly above chance. (b) For the multi-item displays, the earliest detectable color information had a latency of 143 ms in dIPFC and 108 ms in LIP.

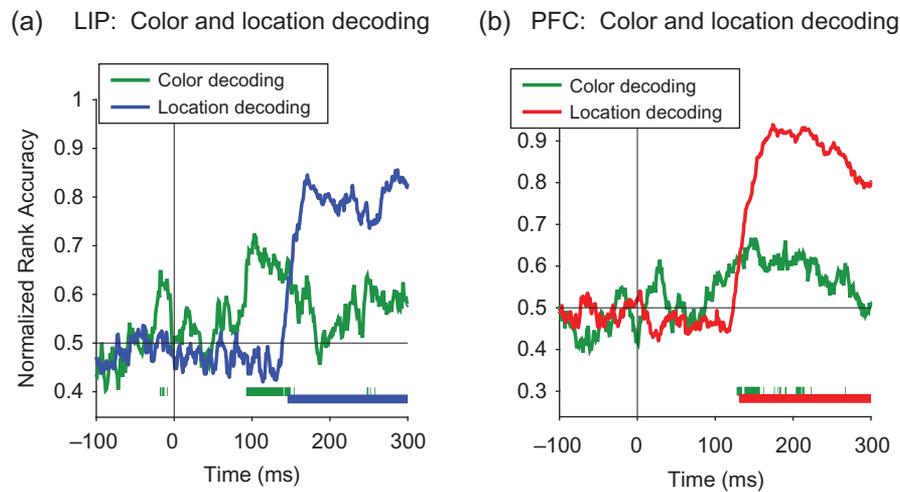


Figure 8. Comparison of the color decoding to spatial location decoding. A comparison of color decoding accuracy and spatial location decoding on the multi-item trials for LIP (a) and PFC (b) using 30 ms bins. Solid bars at the bottom of the plots show when the decoding accuracies for color (green) and location (blue for LIP, red for PFC) are above chance.

LIP and 146 in PFC. Thus it is possible that LIP and dlPFC is involved in the computation converting this low-level color information into a more abstract spatial representation that is needed to resolve the spatial location of the cue.

Discussion

In this article, we examined the time course of firing rate and information changes in LIP and dlPFC while monkeys viewed isolated cue and multi-item pop-out display images. The results show that firing rates first increase in LIP and then in dlPFC for both types of displays (Fig. S1). For isolated cue displays, information about the location of the cue also appears first in LIP and then in dlPFC and the information time course matches the firing rate time course (Figs 2a and 4a, b). Importantly, spatial information for multi-item displays is substantially delayed relative to the firing rate time course, and it appears in dlPFC before LIP (Figs 2b, 3 and 4c, d). This difference between the information about the single-object and multi-item displays is present even though spatial information is largely contained in a stationary code, and represented the same way for both cue colors (Figs 5 and 6). Finally, we see that color information could be decoded from the multi-item displays trials, although overall the color information was relatively weak (Figs 7 and 8). As discussed below, the results shed light on the nature of information processing in the prefrontal and posterior parietal cortex, and give new insights into how pop-out visual displays are processed.

Representation of Pop-Out Information in the Cortex

The established view of pop-out processing posits that pop-out elements are detected by a feed-forward sweep of information along the dorsal pathway, with location information first in the parietal and then in prefrontal cortex, and a common mechanism to operate for single-object and multi-item displays (Buschman and Miller 2007; Siegel et al. 2015). Recent evidence, however, suggests this dorsal stream feed-forward account might be too simplistic (Awh et al. 2012; Khorsand et al. 2015). Previous work by Katsuki and Constantinidis (2012) analyzed

data from multi-item pop-out display trials and found that single-neuron responses in dlPFC contained location information no later than LIP. Here we used population decoding analyses to extend these findings and show that for the multi-item display trials, information appears first in dlPFC relative to LIP which is inconsistent with the simple dorsal stream feed-forward account. A similar pattern of results of information being present in “higher brain regions” prior to “lower level areas” has also been seen by researchers who compared activity in FEF and LIP (Schall et al. 2007; Cohen et al. 2009; Monosov et al. 2010; Zhou and Desimone 2011; Gregoriou et al. 2012; Purcell et al. 2013; Pooresmaeili et al. 2014).

The fact that our results show information first in dlPFC prior to LIP for the multi-item displays while the study of Buschman and Miller (2007) found the opposite temporal pattern raises the question of why do the results differ. We believe that the key difference is that in our study the color of the pop-out stimulus was unknown to the monkey prior to its appearance, while in the study of Buschman and Miller (2007) the monkeys were cued ahead to time of the color of the pop-out stimulus. By cueing the monkeys ahead of time, it is possible that color-selective input that matched the distractor color was suppressed by anticipatory processing, so that information about the pop-out stimulus’ location was available in the first sweep of neural activity; this is the essence of top-down control (Katsuki and Constantinidis 2014).

The idea that anticipatory processing could filter out a particular color in advance is consistent with the findings of Bichot et al. (1996) who showed that spatial information about a pop-out stimulus is present in the first initial increase in firing rates in FEF when monkeys have been over-trained on displays where the distractors and cue always had the same colors on all trials. Indeed, the findings of Bichot et al. (1996) appear similar to the fact that location information was present in the first wave of neural activity during the isolated cue trials, and suggests that similar neural processing might be involved in processing predictable displays as is involved in processing isolated cue displays. Further support for this type of anticipatory filtering comes from psychophysics findings that have shown reaction times are 50–125 ms faster when humans need

to resolve a spatial location from predictable pop-out displays compared with when the color of the pop-out stimulus is not known prior to the start of a trial (Bravo and Nakayama 1992; Wolfe et al. 2003; Leonard and Egeth 2008). This additional 50–125 ms of processing time seen in psychophysics experiments for unpredictable displays appears very similar to the additional 67–94 ms processing time needed to resolve the spatial location in multi-item display trials (Fig. 3 and Table 1) and further suggests that the neural processing we observed in this study might underlie the behavioral effects observed in humans.

Potential Pathways Involved in Converting Color Information to Spatial Information

The fact that color information needs to be converted into spatial information on the multi-item displays raises the question of which brain regions are central to this conversion process. Given that color information was present in LIP and dlPFC prior to spatial information (Figs 7 and 8), one possibility is that the color to spatial conversion is occurring in these brain regions, although this explanation does not account for the facts that spatial information was first in dlPFC prior to LIP, that color information appears relatively weak in these regions (Fig. 8), and that spatial information is invariant to the color of the cue as soon as spatial information is present in these brain regions (Fig. 6). An alternative possibility, outlined in Figure 9, is that the color to spatial information conversion is occurring in the ventral pathway and then information is passed on to dlPFC and then on to LIP (also see Koch and Ullman 1985; Chikkerur et al. 2010; Borji and Itti 2013 which contain similar ideas). In this scenario, visual information travels through both the dorsal and a ventral pathway (Mishkin et al. 1983), where information in the dorsal pathway propagates from low-level areas to LIP and finally to dlPFC, while in the ventral pathway, information propagates from lower level visual areas to color-selective midlevel visual areas that convert color information to spatial information, then on to dlPFC and finally to LIP.

To see how this second scenario (shown in Fig. 9) is consistent with the results in this article, we observe that for the isolated cue trials, both firing rate latencies (Fig. S1a, S1b) and information latencies (Fig. 2a, b) are earlier in LIP than in dlPFC, which is consistent with information flowing in a feed-forward manner through the dorsal pathway. Conversely, for the multi-item display trials information about the cue location travels

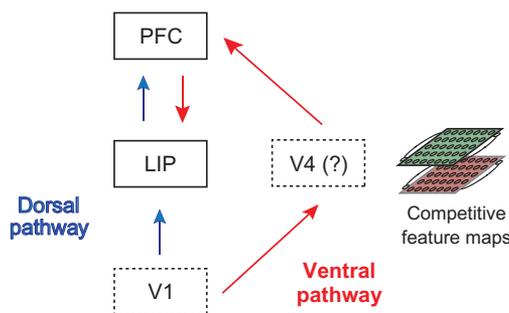


Figure 9. A 2-pathway diagram for how spatial information travels through the dorsal and ventral visual pathways when processing pop-out displays. For the isolated cue displays, information about the location of the cue primarily travels along the dorsal pathway from V1 to LIP to PFC (blue arrows). For the multi-item displays, information about the cue location is resolved in the competitive/contrastive feature maps in the ventral pathway (most likely in V4) before the spatial information travels to PFC and into LIP (red arrows).

through the ventral pathway. When the multi-item displays are shown, firing rates latencies still increase in LIP before dlPFC due to input from the dorsal pathway (Fig. S1c, S1d), however, these initial firing rate increases occur in all neurons, and thus there is no information about the location of the cue (Figs 4c, d); the early color information in LIP and dlPFC is present because these areas are connected to color-selective regions but LIP and dlPFC are not actively using this information which could explain why the information is relatively weak and transient in these regions. Information also simultaneously flows through the ventral pathway, to visual areas such as area V4, where the location of the cue stimulus is selectively enhanced via competitive/contrastive interactions in color maps (Koch and Ullman 1985; Itti and Koch 2001). Information about the cue location then travels to dlPFC, and then propagates to LIP. The idea of processing occurring in 2 pathways suggests that there are 2 waves of input into LIP and dlPFC for the multi-item displays, where the first wave comes from the dorsal pathway and second one coming from the ventral pathway. In fact, these waves of firing rate increases can be seen as 2 peaks in the firing rate plots when the data is plotted comparing trials when the cue is in a neuron's receptive field to trials when the cue is not in a neuron's receptive field (Katsuki and Constantinidis 2012, Fig. S6a, S6b), and has also been seen in previous work (Bisley et al. 2004). In contrast, only 1 peak of firing rate increases is seen for the isolated cue displays, which is primarily driven by the input coming from the dorsal pathway (Fig. S6c, S6d).

Results in the literature are also consistent with information travel over 2 pathways. Of particular relevance is a study by Schiller and Lee (1991) in which monkeys needed to detect an odd-ball stimulus in an array of stimuli that could either be more or less salient (e.g., more or less bright) than other stimuli in a display. Intact monkeys were able to perform proficiently at this task, however, after lesioning V4, monkeys were only able to perform proficiently when the odd-ball stimulus was more salient than the surrounding stimuli. Figure 9 explains this counterintuitive result based on the fact that when a more salient stimulus is shown, a strong drive could carry spatial location over the dorsal pathway leading to intact spatial representations in LIP and dlPFC. In contrast, when a less salient odd-ball stimulus is shown, spatially selective information will not be present in the dorsal pathway (since all locations except the location of the odd-ball stimulus will have high firing rates), and lesioning V4 would prevent the enhancement of the weaker drive in the odd-ball stimulus—thus spatial representations of the odd-ball stimulus will not form in dlPFC and LIP. Recordings made by Ibos et al. (2013) show that when a more salient stimulus is shown, information is present in LIP prior to FEF while when a less salient stimulus is shown FEF is active prior to LIP, which is consistent with our proposal. Further findings in the literature are also consistent, such as the fact that lesioning dlPFC causes a decreased ability to detection of pop-out stimuli in multi-item displays (Iba and Sawaguchi 2003; Rossi et al. 2007), which can be explained by the fact that the default pathway from dlPFC to LIP has been disrupted and information must travel over other neural pathways (this same pathway might also account of the transitory hemineglect seen with PFC damage). Additionally, a transcranial magnetic stimulation (TMS) study that deactivated the inferior parietal lobule in humans showed that there are 2 time windows in which TMS disrupts performance and led the authors to similar conclusions that spatial information travels along 2 neural pathways (Chambers et al. 2004).

Conclusions

In this article, we report several new findings about how the visual system processes pop-out displays. Rather than describing attention in cognitive terms, such as “bottom-up” and “top-down,” we offer a diagram that can explain how attentional demands influence dynamically different parts of the visual system, depending on the incoming information. These results and diagram shown in Figure 9 help clarify that top-down and bottom-up attention do not constitute an adequate dichotomy for describing visual attention generally, across different conditions (Awh et al. 2012). This is useful for both interpreting the large body of existing behavioral and electrophysiological results and for generating testable predictions in future studies which can help either verify or refute the flow of information proposed in this article. We anticipate that turning this conceptual diagram into a working computational model will be beneficial in generating additional insights in the future.

Supplementary Material

Supplementary data are available at *Cerebral Cortex* online.

Funding

Center for Brains, Minds and Machines, funded by The National Science Foundation (NSF) STC award (CCF-1231216) to E.M.; and by the National Eye Institute of the National Institutes of Health under award number (R01 EY016773) to C.C.

Notes

Additional support comes from Adobe, HondaResearch Institute USA, and a King Abdullah University Science and Technology grant to B. DeVore. We wish to thank Kathini Palaninathan for technical help, and Tomaso Poggio for his continual support. Additional supplemental material can be found on figshare at: <https://figshare.com/s/4fe326d1d3a3a94e7de9>. Conflict of Interest: None declared.

References

- Awh E, Belopolsky AV, Theeuwes J. 2012. Top-down versus bottom-up attentional control: a failed theoretical dichotomy. *Trends Cogn Sci*. 16:437–443.
- Bichot NP, Schall JD, Thompson KG. 1996. Visual feature selectivity in frontal eye fields induced by experience in mature macaques. *Nature*. 381:697–699.
- Bisley JW, Krishna BS, Goldberg ME. 2004. A rapid and precise on-response in posterior parietal cortex. *J Neurosci*. 24:1833–1838.
- Borji A, Itti L. 2013. State-of-the-art in visual attention modeling. *IEEE Trans Pattern Anal Mach Intell*. 35:185–207.
- Bravo MJ, Nakayama K. 1992. The role of attention in different visual-search tasks. *Percept Psychophys*. 51:465–472.
- Buschman TJ, Miller EK. 2007. Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science*. 315(5820):1860–1862.
- Cavada C, Goldman-Rakic PS. 1989. Posterior parietal cortex in rhesus monkey: II. Evidence for segregated corticocortical networks linking sensory and limbic areas with the frontal lobe. *J Comp Neurol*. 287:422–445.
- Chambers CD, Payne JM, Stokes MG, Mattingley JB. 2004. Fast and slow parietal pathways mediate spatial attention. *Nat Neurosci*. 7:217–218.
- Chikkerur SS, Serre T, Tan C, Poggio T. 2010. What and where: a Bayesian inference theory of attention. *Vision Res*. 50(22):2233–2247.
- Cohen JY, Heitz RP, Schall JD, Woodman GF. 2009. On the origin of event-related potentials indexing covert attentional selection during visual search. *J Neurophysiol*. 102:2375–2386.
- Connor CE, Egeth HE, Yantis S. 2004. Visual attention: bottom-up versus top-down. *Curr Biol*. 14:R850–R852.
- Constantinidis C, Steinmetz MA. 2001. Neuronal responses in area 7a to multiple-stimulus displays: I. Neurons encode the location of the salient stimulus. *Cereb Cortex*. 11:581–591.
- Corbetta M, Shulman GL. 2002. Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci*. 3:215–229.
- Duncan J, Humphreys GW. 1989. Visual search and stimulus similarity. *Psychol Rev*. 96:433–458.
- Felleman D, Van Essen D. 1991. Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex*. 1(1):1–47.
- Goldberg ME, Gottlieb JP, Kusunoki M. 1998. The representation of visual salience in monkey parietal cortex. *Nature*. 391:481–484.
- Gregoriou GG, Gotts SJ, Desimone R. 2012. Cell-type-specific synchronization of neural activity in FEF with V4 during attention. *Neuron*. 73:581–594.
- Heller J, Hertz JA, Kjær TW, Richmond BJ. 1995. Information flow and temporal coding in primate pattern vision. *J Comput Neurosci*. 2:175–193.
- Hung CP, Kreiman G, Poggio T, DiCarlo JJ. 2005. Fast readout of object identity from macaque inferior temporal cortex. *Science*. 310(5749):863–866.
- Iba M, Sawaguchi T. 2003. Involvement of the dorsolateral prefrontal cortex of monkeys in visuospatial target selection. *J Neurophysiol*. 89:587–599.
- Ibos G, Duhamel J-R, Ben Hamed S. 2013. A functional hierarchy within the parietofrontal network in stimulus selection and attention control. *J Neurosci*. 33:8359–8369.
- Itti L, Koch C. 2001. Computational modelling of visual attention. *Nat Rev Neurosci*. 2:194–203.
- Kaas JH. 2012. The evolution of neocortex in primates. *Prog Brain Res*. 195:91–102.
- Karlen SJ, Krubitzer L. 2007. The functional and anatomical organization of marsupial neocortex: evidence for parallel evolution across mammals. *Prog Neurobiol*. 82:122–141.
- Katsuki F, Constantinidis C. 2012. Early involvement of prefrontal cortex in visual bottom-up attention. *Nat Neurosci*. 15:1160–1166.
- Katsuki F, Constantinidis C. 2014. Bottom-up and top-down attention: different processes and overlapping neural systems. *Neuroscientist*. 20:509–521.
- Khorsand P, Moore T, Soltani A. 2015. Combined contributions of feedforward and feedback inputs to bottom-up attention. *Front Psychol*. 6:155.
- King J-R, Dehaene S. 2014. Characterizing the dynamics of mental representations: the temporal generalization method. *Trends Cogn Sci*. 18:203–210.
- Koch C, Ullman S. 1985. Shifts in selective visual attention: towards the underlying neural circuitry. *Hum Neurobiol*. 4:219–227.
- Leonard CJ, Egeth HE. 2008. Attentional guidance in singleton search: an examination of top-down, bottom-up, and intertrial factors. *Vis Cogn*. 16:1078–1091.
- Meyers EM. 2013. The neural decoding toolbox. *Front Neuroinform*. 7:8.
- Meyers EM, Borzello M, Freiwald WA, Tsao D. 2015. Intelligent information loss: the coding of facial identity, head pose,

- and non-face information in the macaque face patch system. *J Neurosci.* 35:7069–7081.
- Meyers EM, Freedman DJ, Kreiman G, Miller EK, Poggio T. 2008. Dynamic population coding of category information in inferior temporal and prefrontal cortex. *J Neurophysiol.* 100: 1407–1419.
- Meyers EM, Kreiman G. 2012. Tutorial on pattern classification in cell recording. In: Kriegeskorte N, Kreiman G, editors. *Visual population codes*. Boston: MIT Press. p. 517–538.
- Meyers EM, Qi X-LL, Constantinidis C. 2012. Incorporation of new information into prefrontal cortical activity after learning working memory tasks. *Proc Natl Acad Sci.* 109: 4651–4656.
- Mishkin M, Ungerleider LG, Macko KA. 1983. Object vision and spatial vision: two cortical pathways. *Trends Neurosci.* 6: 414–417.
- Mitchell TM, Hutchinson R, Niculescu RS, Pereira F, Wang X, Just M, Newman S. 2004. Learning to decode cognitive states from brain images. *Mach Learn.* 57:145–175.
- Monosov IE, Sheinberg DL, Thompson KG. 2010. Paired neuron recordings in the prefrontal and inferotemporal cortices reveal that spatial selection precedes object identification during visual search. *Proc Natl Acad Sci U S A.* 107: 13105–13110.
- Murray JD, Bernacchia A, Roy NA, Constantinidis C, Romo R, Wang X-J. 2017. Stable population coding for working memory coexists with heterogeneous neural dynamics in prefrontal cortex. *Proc Natl Acad Sci U S A.* 114:394–399.
- Nowak LG, Bullier J (1997) The timing of information transfer in the visual system. In: Rockland KS, Kaas JH, Peters A, editors. *Cerebral cortex volume 12: extrastriate cortex of primates*. New York: Plenum Press. pp. 205–241.
- Pooresmaeili A, Poort J, Roelfsema PR. 2014. Simultaneous selection by object-based attention in visual and frontal cortex. *Proc Natl Acad Sci U S A.* 111:6467–6472.
- Purcell BA, Schall JD, Woodman GF. 2013. On the origin of event-related potentials indexing covert attentional selection during visual search: timing of selection by macaque frontal eye field and event-related potentials during pop-out search. *J Neurophysiol.* 109:557–569.
- Quiñ Quiroga R, Panzeri S. 2009. Extracting information from neuronal populations: information theory and decoding approaches. *Nat Rev Neurosci.* 10:173–185.
- Rossi AF, Bichot NP, Desimone R, Ungerleider LG. 2007. Top down attentional deficits in macaques with lesions of lateral prefrontal cortex. *J Neurosci.* 27:11306–11314.
- Schall JD, Hanes DP. 1993. Neural basis of saccade target selection in frontal eye field during visual search. *Nature.* 366:467–469.
- Schall JD, Pare M, Woodman GF. 2007. Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science.* 318:44–44.
- Schiller PH, Lee K. 1991. The role of the primate extrastriate area V4 in vision. *Science.* 251:1251–1253.
- Siegel M, Buschman TJ, Miller EK. 2015. Cortical information flow during flexible sensorimotor decisions. *Science.* 348: 1352–1355.
- Stokes MG, Kusunoki M, Sigala N, Nili H, Gaffan D, Duncan J. 2013. Dynamic coding for cognitive control in prefrontal cortex. *Neuron.* 78:364–375.
- Treisman AM, Gelade G. 1980. A feature-integration theory of attention. *Cogn Psychol.* 12:97–136.
- Wolfe JM, Butcher SJ, Lee C, Hyle M. 2003. Changing your mind: on the contributions of top-down and bottom-up guidance in visual search for feature singletons. *J Exp Psychol Hum Percept Perform.* 29:483–502.
- Wolfe JM, Horowitz TS. 2004. Opinion: what attributes guide the deployment of visual attention and how do they do it? *Nat Rev Neurosci.* 5:495–501.
- Zhang Y, Meyers EM, Bichot NP, Serre T, Poggio TA, Desimone R. 2011. Object decoding with attention in inferior temporal cortex. *Proc Natl Acad Sci.* 108:8850–8855.
- Zhou H, Desimone R. 2011. Feature-based attention in the frontal eye field and area V4 during visual search. *Neuron.* 70: 1205–1217.