Feature-based attention prioritizes the processing of non-spatial features across the visual field. Classical studies revealed a feature-similarity gain modulation of sensory neuron’s activity. While early studies that quantified behavioral performance have provided support for this model, recent studies have revealed a non-monotonic, surround suppression effect in near feature space. The attentional suppression effects may give rise to a highly limited capacity when selecting multiple features, as documented by studies manipulating the number of attended features. These effects of feature-based attention are likely due to attentional control mechanisms exerting top-down modulations, which have been linked to neural signals in the dorsal frontoparietal network. The neural representation of attentional priority at multiple levels of the visual hierarchy thus shape visual perception and behavioral performance.

Address
Department of Psychology, Michigan State University, East Lansing, MI, United States

Corresponding author: Liu, Taosheng (tsliu@msu.edu)

Visual perception is highly selective. The most prominent and well-understood selection mechanism is spatial attention, where sensory input from a restricted part of visual field attains prioritized processing. Indeed, overt attention with eye movement is intrinsically spatial. However, attention can also be directed to non-spatial visual attributes, such as features and objects [1]. The role of features in visual attention is epitomized in visual search, one of the most studied tasks in visual attention, in which observers need to find a pre-specified target in a visual array (Figure 1a). This task requires using feature information to guide spatial attention and thus search performance is likely influenced by a mixture of both spatial and feature attention. For this reason, studies of feature-based attention often resort to a strategy to fix the locus of spatial attention, by either presenting multiple features simultaneously in the same spatial location (Figure 1b) or presenting probes away from the spatially attended location (Figure 1c). In these paradigms, the focus of spatial attention is kept constant while the attended feature is manipulated (e.g. attend to red versus green color), thus allowing researchers to isolate pure feature-specific selection. These paradigms have helped establish classical findings in feature-based attention (see Box 1), culminating in the feature-similarity gain model: a visual neuron’s response is monotonically modulated as a function of the similarity between the attended feature and the neuron’s preferred feature and such modulation can spread globally. The present review provides a focused update of the most recent results regarding the functional characteristics and physiological underpinnings of feature-based attention. For more expanded treatments, the reader is referred to excellent reviews of the earlier literature [2,3*].

How does feature-based attention modulate feature-selective processing?

The central tenet of the feature-similarity gain model stipulates that attentional modulation of neuronal response is monotonically related to the similarity between attended and preferred feature. It is worth pointing out that this observation was observed in one cortical area (MT) for one feature dimension (motion) under a particular stimulus/task protocol. Thus, it is important to know whether this model describes a general principle of feature-based attention. In particular, it is useful to examine the behavioral implications of the feature-similarity gain model. A number of psychophysical studies have used various techniques to measure the behavioral tuning profile of attentional modulation in a variety of feature dimensions [14–17]. These studies have generally found a monotonic performance modulation, with an enhanced processing of the attended feature and gradual decline of enhancement into suppression for very different features (Figure 2a). Such results generally support the feature-similarity gain model, as monotonic neural modulation should give rise to monotonic performance modulation.

However, a recent study showed that feature-based attention could also modulate performance in a non-monotonic fashion [18*]. Participants were cued to attend to a color during a detection task, while the target’s color was systematically varied in the vicinity of the cued color in a standard color space. The highest detection accuracy was observed for the cued color, but a non-monotonic performance pattern was observed for the other colors (Figure 2b). This ‘Mexican-hat’ performance modulation
Experimental paradigms to study feature-based attention, with example stimulus displays. (a) A stimulus display in a visual search experiment. The task is to find the letter ‘T’ among rotated T’s and L’s. (b) A stimulus display in a feature-based attention experiment, showing two spatially superimposed sets of colored dots. Participants can be instructed to attend to a specific color. (c) Another stimulus display used in feature-based attention experiment. Participants are instructed to attend to either the upward and downward motion in the right stimulus and ignore the left stimulus, while perceptual and neural effects of the left stimulus are measured. This protocol is often used to demonstrate the global spread of feature-based attention.

is interpreted as a surround suppression effect that helps the isolation of the attended feature, analogous to such effects in the spatial domain [19]. This effect is clearly different from predictions of the feature-similarity gain model and previous empirical findings. One important difference between these studies is that previous studies supporting feature-similarity gain usually tested large feature offsets, whereas this study reporting surround suppression tested much smaller offset values in a narrow range. Thus, it is possible that both feature-similarity gain and surround suppression are at work, but at different similarity scales. Consistent with this conjecture, a recent study sampled a relatively large range of colors at a fine resolution and found both a Mexican hat profile at small color offsets and further suppression at large color offsets [20**]. Intriguingly, this study also revealed that the suppressive surround coincided with the color category boundary. Thus, for color, a function of surround suppression might be to improve the distinctiveness of the attended color category. The implication of this category-based effect for other feature domains such as orientation and motion direction remains to be investigated.

Using a frequency-tagging method, the same previous study also observed reduced SSVEP/EEG response to the suppressed color [18*], demonstrating a surround suppression effect in gross measures of neural activity. However, the origin and precise neuronal basis of surround suppression remain unclear. Indeed, the feature-similarity gain modulation observed in the classical studies [5,6] is seemingly incompatible with a surround suppression effect. Interestingly, some neurophysiological studies have reported shifts in neuronal tuning toward the attended feature under certain conditions [21,22], and model simulations have shown that these shifts could lead to surround suppression in behavioral performance [20**]. More work is needed to fully characterize the effect of feature-based attention on behavior and neural population responses.

Capacity of feature-based attention
Attention is widely believed to be an adaptive mechanism that copes with the brain’s limited capacity via selective processing. Yet, the selection itself may also be subject to certain capacity limits. Selecting one task-relevant feature has an obvious advantage compared to either selecting none or selecting the wrong feature. However, it is less clear what occurs when observers try to select multiple features.
Previous studies in visual search have provided somewhat conflicting results. On the one hand, studies have found that searching for two targets is less efficient than searching for one target [23,24], suggesting a limit in attending to two features. On the other hand, other studies have claimed that participants can maintain at least two active attentional templates [25–27]. These studies used very different experimental paradigms and examined different aspects of performance (e.g. search time versus distracter cost), making it difficult to directly compare them. The complexity of visual search tasks (see earlier discussion) also makes it hard to ascertain the processing stages that give rise to the limit [28].

A number of studies have investigated the capacity of feature-based attention in non-search tasks. Indeed, earlier studies have shown that attending to two different features (e.g. upward and downward motion) are more difficult than attending to a single feature when these features are simultaneously presented in different locations [29,30]. In these studies, subjects had to split their spatial attention to two locations and often need to monitor the features over extended time periods. Thus performance limits could arise from an impairment in simultaneously processing two different features, as opposed to selecting [30]. To address these concerns, a number of studies employed a feature detection task in noise, while varying the number of pre-cued features to direct participants’ attention to either one or two features [31*,32]. Compared to a baseline without pre-cue, attending to one feature improved performance, whereas attending to two features also improved performance, but at a significantly lower level than attending to one feature (Figure 2c). In these tasks, only a single target was presented on each trial, thus performance limit is likely due to selection, as opposed to processing multiple targets. Interestingly, the performance benefit for two-cue trials can be predicted by a simple model where participants simply picked one feature to attend on these trials, such that their performance was a mixture of one-cue and no-cue conditions. This observation suggests that participants can only attend to a single feature at a time—a very severe selection limit. Such a limit might be due to an inability to enhance multiple features if attending to each feature evokes large zones of suppression, due to both feature-similarity gain and surround suppression. In addition, because working memory is believed to contain representations that bias attentional selection [33], the observed selection limit is also consistent with suggestions that working memory contains a single active item in the focus of internal attention [34,35]. Further research is needed to understand the locus of this attentional limit, at both the functional and physiological level.

**Neural mechanisms of feature-based attentional control**

So far this discussion of feature-based attention has focused on its effect in perceptual performance and neural responses in early visual system. The fact that
attentional selection can be driven by task goals (e.g. by instructing participants to attend to a specific feature) raises an obvious, and important, question: what are the neural mechanisms that control feature-based selection?

It is useful to briefly consider the analogous question in the spatial domain, where it is relatively well established that a dorsal frontoparietal network (FPN) controls spatial attention. Early studies revealed preparatory neural activity in these areas linked to spatial cues, demonstrating the endogenous nature of such signals [36]. Later studies provided more mechanistic insights regarding the representational basis of these neural signals. A key discovery is the finding that FPN areas contain topographic maps of visual space [37]. These spatiotopic maps thus provide a natural substrate to represent spatial priority, for example, by highlighting locations on these high-level maps and sending spatially specific feedback to sensory areas [38]. Anatomical evidence supporting this scheme has been found in white matter connection patterns from posterior parietal areas to visual areas [39]. Functional evidence has been found in microstimulation studies in monkeys and TMS studies in humans, which have shown a causal link between frontoparietal neural activity and spatial attention [40,41].

These observations and concepts have informed the studies of feature-based attention. Given the evidence that spatial priority is represented by spatiotopic signals in high-level areas, feature-specific signals likely also exist in high-level brain areas that represent featural priority. Several studies have used the two superimposed feature paradigm (see Figure 1b) and compared neural activity when participants attended to either feature [42,43]. In this design, because the physical stimulus is kept constant, any observed neural differences necessarily reflect the modulation due to attention. Perhaps not surprisingly, overall fMRI BOLD amplitude did not show difference when attention was directed to either feature, as these features (e.g. leftward versus rightward motion, red versus green color) should be equivalent a priori. However, a classifier that extracts multivoxel patterns arising from small differences between conditions [44] were able to reliably decode the attended feature. Importantly, these pattern differences were observed in the FPN, including intraparietal sulcus and areas along the precentral sulcus (Figure 3a & b). Thus, neural signals in these areas possess featural selectivity, which qualifies them as candidates for representing priority. The contribution of FPN to feature-based attention was further examined in a more recent study by relating FPN activity to behavioral performance. Participants in this study performed a difficult detection task that required selection of a specific motion direction [45**]. In an fMRI experiment, it was observed that neural patterns in FPN for attending to different directions were more discriminable on correct than on incorrect trials. A follow-up TMS study targeted one FPN area, the posterior IPS, and found that neurodisruption impaired behavioral performance on a task requiring feature selection but did not alter performance on a control task that did not require feature selection (Figure 3c). These results demonstrate the behavioral relevance of FPN neural activity and their causal role in directing feature-based attention.

A further question concerns the relationship and potential functional differentiation among FPN areas, which has not been extensively studied in the realm of feature-based attention. A recent fMRI study investigated network effective connectivity underlying global feature-based attention (Figure 1c), using dynamic causal modeling and Granger causality modeling [46**]. These methods allow the assessment of the direction of information flow among brain regions. A series of extensive analyses revealed that the inferior frontal junction (IFJ) provided the source of information flow for both parietal and visual
areas. Thus, the IFJ might serve as the ultimate source of control signals for top–down feature-based attention.

Studies of object-based attention are also informative as an object can be defined as a collection of features. Hence, object-based attention likely relies on similar mechanisms as feature-based attention. In a MEG study, participants were instructed to attend to either faces or houses in a superimposed face-house movie. Several frontoparietal areas showed object selectivity \[^{[47]}\]. Critically, the IFJ showed synchronized activity with ventral visual areas (FFA and PPA) and the phase of MEG signals in IFJ preceded those in ventral visual areas. This timing difference thus further supports the notion that IFJ coordinates the selection of visual objects in ventral visual areas. Corroborating evidence has also been found in a single-unit study when monkeys searched for specific objects \[^{[48]}\]. In a further study, participants attended to one of two superimposed objects that dynamically varied in multiple feature dimensions \[^{[49]}\]. Even with this highly abstract definition of object-hood, FPN neural activity patterns were found to discriminate the two different objects when they were attended. These results thus suggest that FPN also plays an important role in controlling object-based selection and neural signals in these areas can represent highly abstract information.

Conclusions

Classical studies of feature-based attention have informed an overarching framework as encapsulated by the feature-similarity gain model. Recent studies have expanded this framework by demonstrating a surround suppression effect in feature space and characterizing the capacity limitations of feature-based selection. These perceptual effects of feature-based attention are likely the consequences of attentional modulations of neural activities along the visual hierarchy that are controlled by the frontoparietal network. How neural modulations give rise to a particular performance profile and the precise role of distinct areas in the frontoparietal network remain some of the key questions for future research.

Conflict of interest statement

Nothing declared.

Acknowledgement

I thank Dr. Mengyuan Gong for her help in making Figure 3 and helpful suggestions on an earlier version of the paper. This work was supported by a grant from the National Institutes of Health (R01EY022727).

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest


By varying the similarity between two colors that participants need to attend, the authors found a surround suppression effect (Mexican hat profile) for attention to colors. In a separate experiment using frequency-tagged EEG, a neural correlate of surround suppression was found in SSVEP amplitude.


Using a color detection paradigm, this study also revealed a surround suppression effect for colors. Moreover, suppressive surround was found to coincide with category boundaries, suggesting a perceptual sharpening effect at the categorical level. Computational modeling further showed that surround suppression can be caused by a shift in neuronal tuning.


Participants were cued to attend to one or two colors to detect a single color signal. Performance on two-cue trials can be predicted by a simple model where participants only attended to a single, randomly chosen, color. This study revealed the number of colors that can be attended simultaneously is highly limited.


Participants in this fMRI study performed a difficult threshold detection task contingent on feature selection (attention), as well as a task without a strong demand on selection (baseline). Neural activity patterns showed a higher level of correlation with those in the baseline condition when participants performed correctly on the task attention, demonstrating a neural correlate of task performance. In separate experiments, TMS to posterior parietal cortex selectively impaired performance on a feature-based attention task, demonstrating its causal influence on feature selection.


These authors measured global feature-based attentional modulation throughout the brain and used sophisticated analytic methods to assess effective connectivity among key cortical areas in the visual and frontoparietal cortices. They obtained converging evidence that the inferior frontal junction is the single source region of attentional modulation as measured by information flow, thus suggesting it could issue control signals to direct feature-based selection in other brain areas.


In an MEG experiment, participants viewed a dynamic face-house movie and were cued to attend to either faces or houses. In this object-based selection paradigm, multiple frontoparietal areas showed differential response when attention was directed to either object. In particular, inferior frontal junction (IFJ) showed synchronized activity with ventral visual areas FFA and PPA. Critically, phase analysis showed that IFJ activity preceded FFA and PPA activity. The timing difference is consistent with IFJ as the source of attentional control signals.
