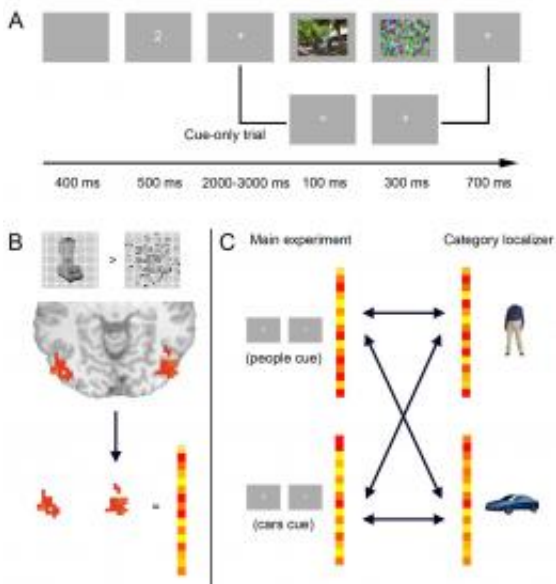# Categories rule: High-order brain centers pave the way for visual recognition

July 11 2011, by Stuart Mason Dambrot

Experimental design and analytical approach. (A) Participants were instructed to detect either people or cars briefly presented in natural scenes preceded by a symbolic cue (a letter or number) indicating the target category. In 33% the cue but no scene was presented, to isolate responses to the cue itself ("cue-only trial"). (B) OSC was localized in each individual participant by contrasting activity to intact versus scrambled objects, presented in a separate experiment. Voxels activated by this contrast were selected for multivoxel pattern analysis. (C) Multivoxel response patterns in the cue-only trials (people and cars cues) in the main experiment were correlated with response patterns evoked by example pictures of people or cars, presented without other visual context in a separate category localizer experiment. The between-category correlations (diagonal comparisons) were subtracted from the within-category correlations (horizontal comparisons) to estimate the category-specificity of cue-related activity. (c) PNAS, doi:10.1073/pnas.1101042108

(Medical Xpress) -- The real world is, in a word, cluttered – but thanks to evolution, we (and other mammals) have no trouble detecting objects in visually complex natural environments. Determining precisely how this occurs is a deceptively complex task, since the retinal and neural mechanisms responsible for simpler percepts – lines, edges and the like –do not account for this survival skill – in fact, they actually interfere with it. Recently, however, scientists have used functional magnetic resonance imaging (fMRI) to elucidate the top-down processes by which high-level cortical areas that deal not with simple percepts, but rather abstract perceptual categories, actually prepare lower-level visual brain centers to perceive detail amidst disorder.

Conducted by Asst. Prof. Marius V. Peelen at the [Center for Mind/Brain Sciences](link) at the University of Trento in Italy with Prof. Sabine Kastner at the Department of Psychology and the [Princeton Neuroscience Institute](link) at Princeton University, the research demonstrates that even when the precise visual characteristics of an object to be found is not known ahead of time, these higher cortical structures mediate visual search.

The research measured activity in a brain area known as the object-selective cortex (OSC) while participants were preparing to find a wide range of representational images of cars or people within briefly-displayed (100 ms) naturalistic scenes which they had not previously viewed. The subjects were first given visual cues that specified the category of objects (i.e., cars or people) to be located within the scenes. The key finding was that the cue alone – that is, even when no scene was subsequently shown – generated OSC responses determined through *multivoxel pattern analysis* (MVPA) that were strikingly similar to those that occurred when looking at actual examples of the cued category. Moreover, when looking at scenes, this neural activity pattern reliably predicted the subjects' performance in detecting the cued visual target.

(Unlike fMRI analysis, which focuses on individual brain voxels (volumetric pixels), MVPA enhances fMRI interpretation by identifying the information in broader patterns of brain activity.)

While the technology used was already established, and so did not present significant challenges, Peelen notes that it takes six seconds to measure a neural signature – so it was needed to overcome the way neural measurements had previously been confounded with visual activity. "We came up with a clever design in which we showed the visual cue without subsequently displaying a scene," he adds. "Since we primary gathered data using this technique, the measured signal reflected brain activity in the absence of visual input."

Given the brain's ability to perceive the world using various senses, and the fact that the research relied on symbolic (rather than visually-specific) cues invoked OSC activity, Peelen says that he expects that his results would be similar with different types of symbolic cues, whether these are spoken or textual. "Indeed, if we search for something in our daily life environment, the trigger to search can come from multiple sources – that is, a thought, but also an external demand – and it is unlikely that the brain has developed different mechanisms for each of these different cues. A very interesting question is how the brain transforms a symbolic cue, such as a word, a thought, or spoken text, to a visual 'search template' that effectively guides visual search. Very little is known about this transformation process."

Peelen notes that one unexpected and interesting result was that activity in the medial prefrontal cortex (MPFC) seems to reflect a high-level source of categories used in preparatory visual search mediation. "An interesting area of follow-up research would be to determine precisely how the MPFC communicates with other visual cortical regions."

In addition to sensory input and internal neocortical activity, the roles of

emotion and memory, and their corresponding brain areas, are intimately involved with perception, attention and motivation – and therefore with preparatory mediation. "It is likely that subcortical structures involved in motivation and arousal play an important role in the temporal aspect of preparation, Peelen reflects. "That is, to successfully perform our task, the participants had to be ready at the moment the scene appeared. However, such temporal preparation would not be expected to be specific to particular object categories, but would operate equally in all cases."

Moreover, he continues, "a general hypothesis that follows from our results is that preparation to detect particular target objects is most effective in brain regions that can discriminate these target objects from distractor objects. In our study, we investigated detection of emotionally neutral object categories. We showed that OSC was best in discriminating these categories and, accordingly, preparatory activity was also the most effective in OSC. However, one could think of situations in which we actively search for emotional – for example, dangerous – objects. In that case, it is in line with our hypothesis that preparatory activity in the amygdala – where emotional or dangerous objects are thought to be discriminated from non-emotional objects – facilitates detection. An alternative theory – in fact, an ongoing debate – is that structures like the amygdala operate independently of top-down control, and will detect emotional stimuli even when one is not actively searching for them."
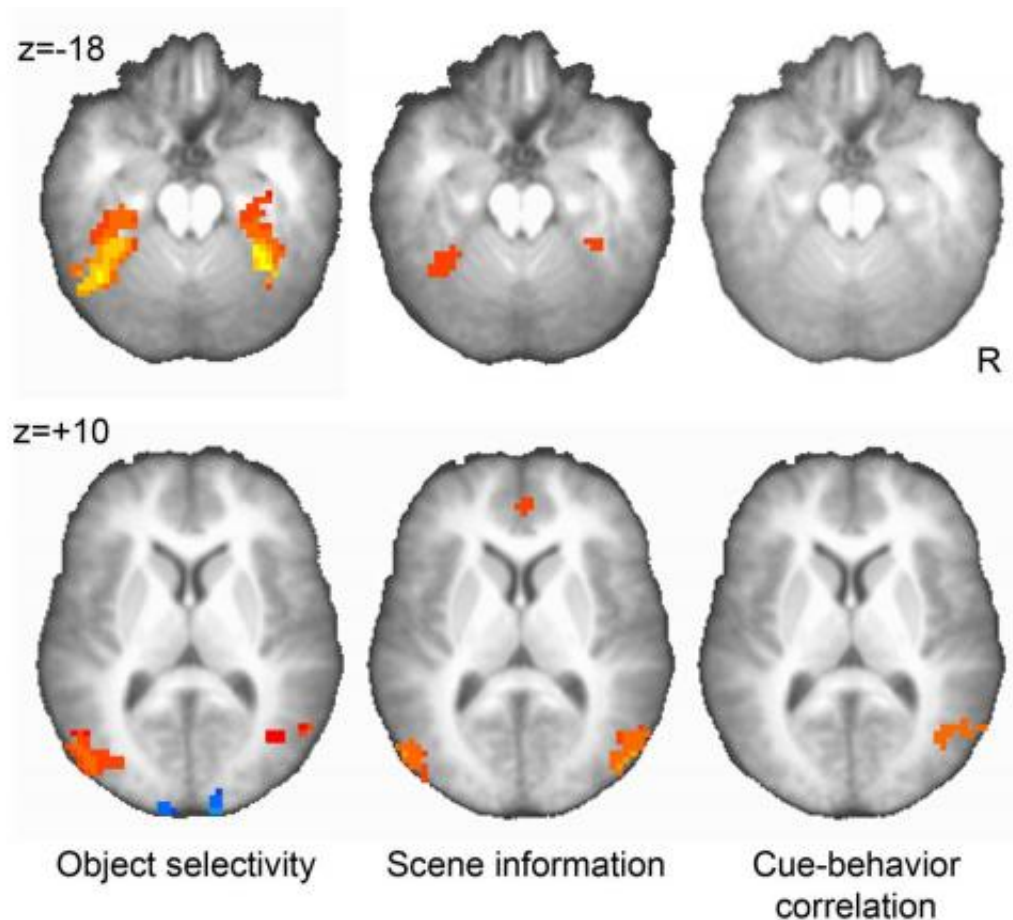
Whole-brain analyses. Results from whole-brain group analyses, overlaid on the group-average anatomical scan. (Left) Ventral and dorsal OSC foci, activated by the univariate contrast between intact and scrambled objects in the OSC localizer. (Center) Occipitotemporal and medial prefrontal clusters from the multivoxel searchlight analysis testing for spheres that discriminated between the two scene types (people vs. cars) on the basis of independent category localizer patterns. (Right) The right occipitotemporal cluster from the multivoxel searchlight analysis testing for spheres with a positive correlation between the category-specific cue effect and behavioral performance. (c) PNAS, doi: 10.1073/pnas.1101042108

In addition, Peelen adds, "memory is of course a broad term, and most forms of knowledge and recognition can be argued to rely in some way

on memory. For example, in our study, participants must have had knowledge of what a person or a car may look like in cluttered scenes to actively prepare themselves for the detection. If such memory of object shape is impaired, it is likely that one won't be able to effectively prepare for object detection, and – in extreme cases – may not even recognize objects as being a person or a car."

Other ways to improve on fMRI scanning might be *optogenetics*, which allows for the controlled switching of individual neurons using brief pulses of light, and *electrical microstimulation*, which uses microelectrode arrays to interface with small groups of neurons with high spatiotemporal precision. "Optogenetics has a potential similar to that of electrical microstimulation," Peelen notes, "although it's thought to be more precise in targeting specific neurons. It could constitute an exciting tool to follow up on our study. For example, our findings showed that preparatory activity in some brain regions is critical for successful object detection. Optogenetics – and perhaps electrical microstimulation as well – could be used to control activity in neurons that code for the target object category."

Of critical importance is that this precise activity could be timed precisely, and could be applied before a visual scene appears. "This would allow us to address several intriguing questions, including the precise time window in which preparatory activity is useful, the specificity of this activity to particular neural populations – for example, those coding for the target category – and perhaps most interestingly, whether externally induced preparatory activity would result in facilitation at all, or whether this needs to be driven by, and activated in concert with, top-down regions such as the MPFC. Indeed, perhaps one could even think of directly stimulating these source regions, and test whether this activity then results in preparatory activity in visual cortex."

In terms of applications, Peelen's initial thoughts look towards novel

research studies, such as working with congenitally blind patients. While such individuals have no visual experience, their visual system's organization is similar to that of sighted individuals – but certain key regions respond to verbal or tactile material. (Braille, for instance, activates the same areas of the visual cortex as does reading in a sighted individual.) "This indicates that the neural activity we see in our research might already be conceptual, rather than visual, in nature," Peelen speculates.

Venturing further afield, Peelen says one potential application may lie in computer vision, such as the automatic labeling of photographs or video by search engines or robots. "Our paradigm may be very well suited for studying the critical object features humans use to perform visual search because it may reveal the object features that are activated during search preparation in the absence of visual input," he observes. "One may then be able to design algorithms that implement these features into computer vision. Considering the automated retrieval of visual information, and specifically the key issue of determining the different analysis pathways that should be used for detecting different semantic categories in photos and video, researchers have made little use of the observation that the human brain is remarkably good in performing these tasks both accurately and rapidly."

Thus, he concludes, "perhaps we should start by looking at the human brain to find inspiration on how automated visual categorization should work. Linking brain science with Information and Communication Technologies by taking advantage of modern brain imaging techniques for the purpose of devising better, cortex-inspired solutions to video search, is a promising research direction."

In terms of future research, Peelen is setting up a study in which real-time fMRI is used to allow participants to view and, hopefully, control their brain activity during visual search preparation. While preparing to

search for particular objects, subjects will be asked to increase activity in either low-level visual areas (which hindered visual search) or high-level visual areas (which facilitated visual search). "Participants will not be aware of which brain region they are controlling on a given day," he explains, "We hope to find that we can manipulate the strategy participants employ, and to improve performance in participants that use a sub-optimal strategy related to low-level visual areas."

Subjects' fMRI scans will be analyzed using fast algorithms to decode neural signatures of object categories and – as was done in the current research – find a match between activity patterns and categories, but in real-time. Peelen notes that this design may allow subjects to be presented with feedback as to how much their brain activity resembles the viewed visual pattern. "For example," he notes, "an auditory stimulus can vary in pitch based on how closely their brain activity matches the category cue. They can then use their own cortical activity to optimize their visual search performance." This is relevant, he adds, because while a given subject's performance is relatively fixed, there are significant differences between subjects.

In addition, he notes that "our finding that different people use different search strategies, as reflected in activation of different brain areas, may have implications for all situations in which visual search/object detection is important, including airport security, military applications, and other areas."

  More information: *A neural basis for real-world visual search in human occipitotemporal cortex*, Published online before print, [doi:10.1073/pnas.1101042108](doi) ; Published *PNAS* July 5, 2011

rewritten or redistributed in whole or part without the express written permission of PhysOrg.com.