

Published in final edited form as:

Trends Cogn Sci. 2009 April ; 13(4): 167–174. doi:10.1016/j.tics.2009.01.008.

Selecting and perceiving multiple visual objects

Yaoda Xu¹ and Marvin M. Chun²

¹Department of Psychology, Harvard University, 33 Kirkland Street, Room 780, Cambridge, MA 02138, USA

²Department of Psychology, Yale University, 2 Hillhouse Avenue, New Haven, CT 06511, USA

Abstract

To explain how multiple visual objects are attended and perceived, we propose that our visual system first selects a fixed number of about four objects from a crowded scene based on their spatial information (object individuation) and then encode their details (object identification). We describe the involvement of the inferior intra-parietal sulcus (IPS) in object individuation and the superior IPS and higher visual areas in object identification. Our neural object-file theory synthesizes and extends existing ideas in visual cognition and is supported by behavioral and neuroimaging results. It provides a better understanding of the role of the different parietal areas in encoding visual objects and can explain various forms of capacity-limited processing in visual cognition such as working memory.

Introduction

Many everyday activities, such as driving on a busy street, require the encoding of multiple distinctive visual objects from crowded scenes. Extending previous behavioral theories and incorporating recent brain imaging and behavioral data, we describe a neural object-file theory to explain how multiple visual objects are attended and encoded. Given processing limitations, our visual system can first select a fixed number of about four objects from a crowded scene, based on their spatial information (object individuation) and then encode their details (object identification). We present evidence showing the involvement of the inferior intra-parietal sulcus (IPS) in object individuation and the superior IPS and higher visual areas in object identification. These two stages of operation could underlie the variety of ways that visual processing is capacity limited, such as in visual short-term memory (also known as visual working memory), enumeration and multiple object tracking.

The neural object-file theory

Our theory consists of two main components, object individuation and object identification (Figure 1).

Object individuation

During object individuation, a fixed number (around four) of visual objects are selected via their spatial locations. Object representations at this stage of processing are coarse and contain only minimal feature information, enough to allow figure-ground segmentation, but not sufficient to support object identification or recognition. Once set up, these 'object file' [1] representations can be updated and elaborated over time during object identification,

when more detailed featural and identity information become available. For example, observers can track a moving object before they are certain of its identity ('It's a bird; it's a plane; it's Superman!' [1]). Object individuation is the initial set-up process of such object files based on object spatial-locations. Our neuroimaging results indicate that object individuation involves the inferior intra-parietal cortex [2–6]. This brain area tracks a fixed number of about four objects, independent of the encoding demand from each object, and takes into account the grouping cues between visual elements in individuating objects. The individuated objects at this stage of processing are not merely abstract location pointers [7], but can be considered as proto objects or abstractions of full-fledged objects [8]. The fixed capacity of object individuation could account for several capacity-limited visual processing findings, as reviewed later, and could explain why visual object perception is typically constrained to the 'magic number four' [9]. Because visual attention can select multiple locations at once without selecting the regions in between [10,11], it is the key object selection mechanism during object individuation and is thus an integral part of our neural object-file theory.

Object identification

During object identification, a subset of the objects selected at the individuation stage is further processed with their detailed featural information as well as the binding between features available to the observers. Superior IPS and the extrastriate visual cortex have been shown to encode objects into great detail [2–6]. The superior IPS could additionally represent feature bindings that define distinctive objects [4]. It is at this stage of processing that observers become aware of object identities, whether they are familiar recognizable objects or novel ones with no prior encounter (e.g. an abstract shape or a unique color and shape conjunction). Here, we have full-fledged objects, corresponding to the content of the object files in Kahneman *et al.* [1]. Object identification has a resolution and resource limited representation. Depending on object complexity, task demand and the representation resolution needed, the number of objects available here is variable. This is reflected both in behavioral measures [2,12] and in the observation that superior-IPS-response tracks the total amount of feature information encoded, independent of the number of objects this information comes from [4]. Thus, more objects can be encoded simultaneously when the encoding demand from each object is low (e.g. a simple object or a simple feature from an otherwise complex object), but fewer can be encoded when the demand from each object is high.

Although object individuation has a fixed capacity limit and represents grouping even when it is task irrelevant, object identification is closely tied to the goals and the intentions of the observer, such that not all object features are always encoded together and to the same precision. For example, object colors might not be encoded if the task only involves object shapes, and the orientation of an object will be encoded with different precision depending on the orientation discrimination needed. The superior IPS region has been shown to represent different object features when they become task relevant [2,13].

Possible computational algorithms

The computational algorithms enabling visual object-selection are well described by saliency map models such as that proposed by Koch and colleagues [14]. Individuation can occur via a winner-takes-all mechanism based on the saliency of low-level visual features. The main difference is that our theory selects four, rather than one, feature-occupied spatial locations (which correspond to potential objects) during individuation. Among the four selected objects, some are likely to be more salient than others (because of a more foveal location, containing a brighter color, etc). This could influence subsequent processing priority for these objects and determines which objects will be encoded when less than four

objects can be represented during object identification. Objects that are individuated but failed to reach the identification stage might still reach awareness as coarsely represented objects. They could guide future fixations and bias processing priority during subsequent object identification to bring the full details of these objects into our awareness.

Implications for parietal mechanisms in visual cognition

Previous studies have highlighted the importance of the parietal cortex in attention-related processing [15]. Our proposed functional subdivision further describes how multiple attentional mechanisms within the parietal cortex can cooperate to select, identify and encode visual objects. Traditionally, it is believed that separate neural mechanisms are involved in localizing objects and in identifying what they are, with such ‘where’ and ‘what’ processing mapped onto distinctive dorsal and ventral visual processing streams, respectively [16]. Although the theoretical distinction between the ‘where’ and the ‘what’ processing remains valid and consistent with our theory, the neural object-file theory describes one way that these two mechanisms can interact. Moreover, this theory argues that the parietal cortex is not only involved in the ‘where’ but also the ‘what’ processing of visual objects, consistent with recent findings [17,18].

Additional characteristics

Because low spatial frequency information of an object is activated earlier than high spatial frequency information during visual processing [19], these two types of information could support object individuation and identification, respectively. Objects from different visual object categories (e.g. faces or letters) are likely to differ in low spatial frequency and those within a category are likely to differ in high spatial frequency. Consequently, information available during object individuation could allow observers to distinguish objects between categories but not within a category [20].

Although object individuation normally precedes object identification, results of object identification can refine or modify initial object individuation. For example, parts of an occluded object might initially be selected as multiple different objects; after shape identification, they can be reselected as comprising a single object [21]. Top down effects such as priming and expectation can also influence what is represented during individuation and identification [22]. Real world object perception, thus, is likely to involve feedback and interactions between object individuation and identification to achieve the most efficient and stable representations for objects [23] (Box 1).

Box 1. Beyond object individuation and identification

The object individuation–identification framework enables the observers to select a limited number of visual objects for subsequent detailed processing. Often in visual perception, we need to zoom in and out of a part of the visual scene, shift visual attention across the different levels of the visual hierarchy and select objects at the appropriate level (i.e. select either the individual trees or the entire forest). To do so, two processing systems are needed: one that tracks the overall hierarchical structure of the visual display, and the other that processes the current objects of attentional selection. The two processing systems can be distinct, or they can be intertwined. Although decisive studies are further needed, our grouping data indicate that the inferior IPS can represent selected objects and the visual hierarchy in which these objects reside. Meanwhile, the LOC and the superior IPS can be parts of the neural mechanism that represents what is most relevant to the current goal of visual processing [3]. Work by Yantis and colleagues [60] indicates that the control of the attentional shift signal originates from the superior parietal lobule, which participates in visual attention shifts between objects, visual

features, spatial locations and even different sensory modalities. The interactions among these different cognitive and neural mechanisms are likely to enable us to perceive both the individual trees and the entire forest.

To comprehend an entire visual scene that usually contains more than a few objects and to overcome the capacity limitations in object individuation and identification, we also need to shift visual attention to different parts of a scene and incorporate information from each snap shot of object individuation–identification to form an integrated scene representation. The parahippocampal place area (PPA) has been shown to participate in scene perception [61]. Further research is needed to understand the interaction between the object individuation–identification network in occipital and parietal cortices and the PPA, and how we perceive a natural scene as an integrated and coherent whole, rather than isolated snapshots.

Evidence supporting the neural object-file theory

The distinction between object individuation and object identification has been noted more than 20 years ago. Sagi and Julesz [24] observed fast target detection but relatively slower target identification performance in a visual search task. Kahneman, *et al.* [1] proposed that spatial and temporal information allows an ‘object file’ to be created (corresponding to object individuation), before they can be filled with object features to allow objects to be identified (corresponding to object identification). Likewise, Pylyshyn [7] argued that a fixed number of approximately four objects are indexed via spatial locations in preattentive visual processing, with featural information available at a later attentive stage of processing. The object individuation–identification distinction has also been noted in infant research [25] (Box 2). Here, we present neural and behavioral evidence to further support the dissociation between object individuation and identification and our neural object-file theory.

Box 2. The development of object individuation and identification

Research in developmental studies also emphasizes the distinction between object individuation and identification, with the development of object individuation preceding object identification [25]. According to some experiments [62], infants are able to use spatiotemporal information to individuate multiple objects at around 5 months, but can not use feature information to do so until much later at ~10 to 12 months. Meanwhile, infants can only retain identity of one hidden object at 6.5 months, and can use shape, but not color, to identify two hidden objects by 9 months [63].

When two different objects were first presented simultaneously before disappearing behind a screen, 10-month-old infants could keep track of both objects; however, when the two objects were shown sequentially one at a time, these infants inferred the presence of only one object [62]. In our fMRI study, when objects were presented sequentially at the same spatial location, whereas superior IPS and LOC continuously tracked VSTM capacity, the inferior IPS response became flat and no longer varied with display set size [2]. This suggests that 10-month-old infants can use the output from the inferior IPS, but not that from the superior IPS and the ventral visual cortex, to infer the number of objects present. It would be exciting in future studies to examine the correspondence between developmental behavioral changes in object perception and the maturation of the parietal areas as described in our theory.

Does the adult human brain exhibit plasticity for object individuation and identification capacity? Action video game playing has been shown to increase performance in several capacity-limited visual processing tasks, such as visual enumeration and object tracking

[64]. Interestingly, although gaming improves enumeration accuracy, it does not increase the subitizing capacity [65]. Visual expertise in a particular object category can result in superior object identification performance and improve VSTM capacity [66,67]. This suggests that object individuation capacity could be fixed once it is developed, whereas that of object identification could be improved by training and experience. Consistent with this view, training and visual expertise induce neural changes corresponding to improved object identification abilities [68–70].

Capacity limitation in object individuation and identification

Visual short-term memory capacity limitation

Behavioral literature disagrees on whether visual short-term memory (VSTM) (also known as visual working memory [VWM]) capacity is limited to a fixed number of objects [26,27], or whether it is resolution-limited and can be flexibly allocated to a variable number of objects depending on object complexity and the encoding demand [12,28–32]. In a recent fMRI study, we examined posterior brain mechanisms supporting VSTM [13,33]. We found that, whereas representations in the inferior IPS were fixed to around four objects or feature occupied spatial locations regardless of object complexity, those in the superior IPS and the lateral occipital complex (LOC) were variable, tracking the number of object shapes held in VSTM, and representing fewer than four objects as their complexity increases [2] (Figure 2). Because results from VSTM (also known as ‘VWM’) tasks can inform the neural mechanisms underlying visual object perception in general [34,35], these results support the dissociation between object individuation and identification. Corresponding to object individuation, the inferior IPS selects a fixed number of approximately four objects via their spatial locations; and corresponding to object identification, the superior IPS and the LOC encode the shape features of a subset of the selected objects into great detail. Thus, the dichotomy in which number and resolution can represent distinct attributes of the VSTM representation, directly map onto the dissociation between object individuation and identification. Because of the fixed four-object limitation in object individuation, the maximal VSTM capacity might not exceed four objects. Meanwhile, because limitations in object identification ultimately determines the number of objects successfully retained in VSTM, final VSTM capacity is variable and depends on object complexity. Nevertheless, with complex items that are difficult to encode (i.e. errors observed even for just one item), VSTM capacity will be determined by the encoding limit during object identification, rather than a selection limit during object individuation [32].

Object-based encoding in the brain

Decades of vision research have shown that objects are discrete units of visual attention and information processing [36–39]. Because object individuation selects distinctive visual objects, the neural object-file theory predicts that grouping between visual elements ought to influence object selection and modulate neural responses in the inferior IPS. Consistent with this prediction, we found that even when grouping cues are task irrelevant, they nevertheless can modulate object individuation in the inferior IPS, resulting in lower response for grouped than for ungrouped visual elements [3,5](Figure 3a). This grouping effect allowed more visual information to be selected and encoded in the object identification stage, resulting in greater information storage in the superior IPS for the grouped visual elements [3] (Figure 3a).

Object-based selection during individuation can account for several behavioral observations showing the importance of object-based representation in visual cognition [36–40]. It can also explain patient’s behavioral performance after parietal brain lesions [40–42]. For example, after unilateral parietal lesions, observers’ ability to perceive the presence of two

objects, one on each side of the space, was greatly improved by connecting the two objects with a bar forming one big object with two parts instead of two separated objects [40].

Balint's syndrome

It has long been known that patients with bilateral parietal lesions can perceive only a single object when confronted with two or more objects, resulting in Balint's syndrome. Meanwhile, these patients can perceive a single complex object, showing a dissociation between intact object identification abilities and impaired object individuation for multiple objects [41,42]. The bilateral parietal lesions that are necessary to cause these deficits coincide well with the bilateral inferior IPS region examined in our studies [2–6], supporting the role of this brain area in object individuation and its necessity in visual perception, such that, without it, visual perception is severely impaired.

Dissociation between individuation and identification

Feature encoding during multiple object tracking

Pylyshyn and Storm [10] reported that observers could simultaneously track about four moving targets among otherwise identical moving distractor objects. Observers' memory for the color or shape features of the successfully tracked objects, however, is surprisingly poor [43]. Even when all objects differed in identity, and identity was present throughout tracking, recall of target location is superior to the recall of target location with identity [44]. How could an observer successfully attend an object continuously without noticing its featural changes or simultaneously encoding its identity? In typical multiple object tracking (MOT) tasks, continuous updating of target spatial locations, but not their identity, is needed. The dissociation between tracking and identification therefore supports the dissociation between object individuation and identification. Our neural object-file theory is able to explain these seemingly puzzling MOT results. Consistent with findings from VSTM studies, the inferior IPS regions have been shown to participate in MOT [45].

Enumeration

In enumeration tasks, observers quickly report the number of objects in a display. The typical finding is that, in the one to four object range, observers can subitize, with accuracy close to perfect and reaction times (RTs) increasing slowly. With more than four objects, accuracy decreases and RTs increase sharply and linearly as observers switch to the slow and laborious counting process [46].

Like others, we believe that subitizing reflects object processing at the individuation stage, with its capacity limitation determining the subitizing capacity. Because object individuation precedes object identification, our neural object-file theory predicts that when enumeration requires detailed object identification, subitizing does not occur. Consistent with this prediction, when the to-be-enumerated items had to be first identified by specific feature conjunctions, subitizing disappeared [47]. Posterior parietal cortex has been shown to exhibit distinctive response patterns for subitizing and counting [48].

A crucial test: neural response to object repetition

Consider a task in which observers are asked to encode the shapes of a set of objects that can all be identical or different from each other and then decide whether a particular object shape is present or absent. According to the neural object-file theory, the inferior IPS should treat these objects as multiple, separate entities in object individuation, regardless of whether they are all identical or all different. By contrast, the superior IPS and higher visual areas should treat multiple identical objects as a single unique object during object identification

because of the same demands on shape feature representation in both cases. This prediction was confirmed in a recent study (Figure 3b), providing independent support for the neural dissociation between object individuation and identification [6]. Importantly, although object identity and not its location was task relevant, object individuation in the inferior IPS was still sensitive to the number of objects present independent of identity, indicating the necessity of individuation in object perception regardless of the need for explicit location encoding.

Discussion

Why does object individuation have a roughly four-item limit? This limitation is likely to have evolved in the visual system as a functional characteristic rather than a deficiency. Pylyshyn [7,49] argued that, to comprehend simple relational geometrical properties, we need to simultaneously reference multiple places or objects in a scene. Having access to four objects or feature occupied spatial locations can be sufficient to comprehend most 3D object locations with respect to each other. After all, in Cartesian space, an object's precise 3D location can be unambiguously determined by its projections on the three orthogonal axes. Including the object itself, that makes four items that need to be simultaneously perceived in the 3D space. Perhaps this is what determines the four-item limit in object individuation. Nevertheless, given that we often encode objects from the same surface (e.g. the ground plane) and that top-down knowledge of the visual world are also available, having simultaneous access to fewer than four items can be sufficient in some cases [33].

Researchers have recently challenged the four-object limit in MOT tasks by showing that more or fewer than four objects can be successfully tracked in certain situations [50]. MOT is a complex and demanding task, involving constant updating of the spatial temporal information of target objects in addition to inhibiting identical-looking distractor objects. Thus, MOT involves both object individuation in the inferior IPS and the control mechanisms in the frontal and/or prefrontal cortex (which could operate in a serial manner). Increasing object moving speed or decreasing object spacing can increase interference from distractors and disrupt the updating process, resulting in fewer than four objects being tracked. In contrast, slowing object speed or increasing object spacing can facilitate chunking and grouping, resulting in more objects being tracked without necessarily increasing object individuation capacity. If we can isolate object individuation from the updating process and prevent grouping, our neural object-file theory would still predict a four-object limit in MOT.

Recent neural imaging studies have reported one or more retinotopic maps in the parietal cortex along the IPS [51–53]. It would be necessary in follow-up studies to examine how the two IPS regions, described in our theory, map onto distinctive parietal retinotopic maps. Object individuation in inferior IPS is likely to have a retinotopic space representation. It remains an open question whether the same applies to the superior IPS during object identification.

It would also be necessary in future studies to apply techniques with excellent temporal resolution such as electroencephalography (EEG) and magnetoencephalography (MEG) to examine the temporal dynamics and the interactions between object individuation and identification. Although the time course of object individuation is likely to be fixed, that of object identification could depend on object complexity and the encoding demand.

Areas around the IPS have also been linked to other cognitive tasks, such as number manipulations beyond simple enumeration [54], decision-related processes [55,56], memory retrieval [57] and multi-sensory integration [58]. Do the same parietal areas involved in the

neural object-file theory participate in these other tasks? Answers to this question would not only provide a better understanding of the parietal cortex in cognition but would also integrate otherwise seemingly unrelated cognitive operations to the same or adjacent brain areas, advancing our understanding on how the brain evolved to perform these sophisticated cognitive operations. It is conceivable that areas originally dedicated to visual processing could later be recruited for non-visual tasks and that the parietal cortex could have a more general role of accumulating, storing and integrating sensory information relevant to behavioral goals and choices [59] (Box 3).

Box 3. Outstanding questions

- How do the inferior and superior IPS regions defined in the neural object-file theory correspond to the retinotopic maps in parietal cortex?
- How do the object individuation and identification systems interact?
- How do we perceive objects in a complex scene when attention is allowed to scan the scene serially and to hierarchically zoom in and out of parts of a scene? How do we go from object perception to scene perception?
- What is the relationship between the parietal cortex's involvement in multiple object representation and its involvement in other higher cognitive operations?

Conclusions

Previous behavioral research and theoretical developments have emphasized the important and useful distinction between object individuation and identification for understanding the selection and encoding of multiple visual objects. Here, we propose how these two processes can be supported by distinct neural mechanisms: object individuation is mediated by the inferior IPS and object identification by higher visual areas and the superior IPS. Our neural object-file theory describes how the visual system selects and perceives multiple objects from complex visual scenes despite limited processing capacity.

Acknowledgments

We thank Nelson Cowan, Edward K. Vogel, Todd D. Horowitz and an anonymous review for comments on early drafts of this paper. This work was supported by NSF grants 0518138 and 0719975 to Y.X. and NIH grant EY014193 to M.M.C.

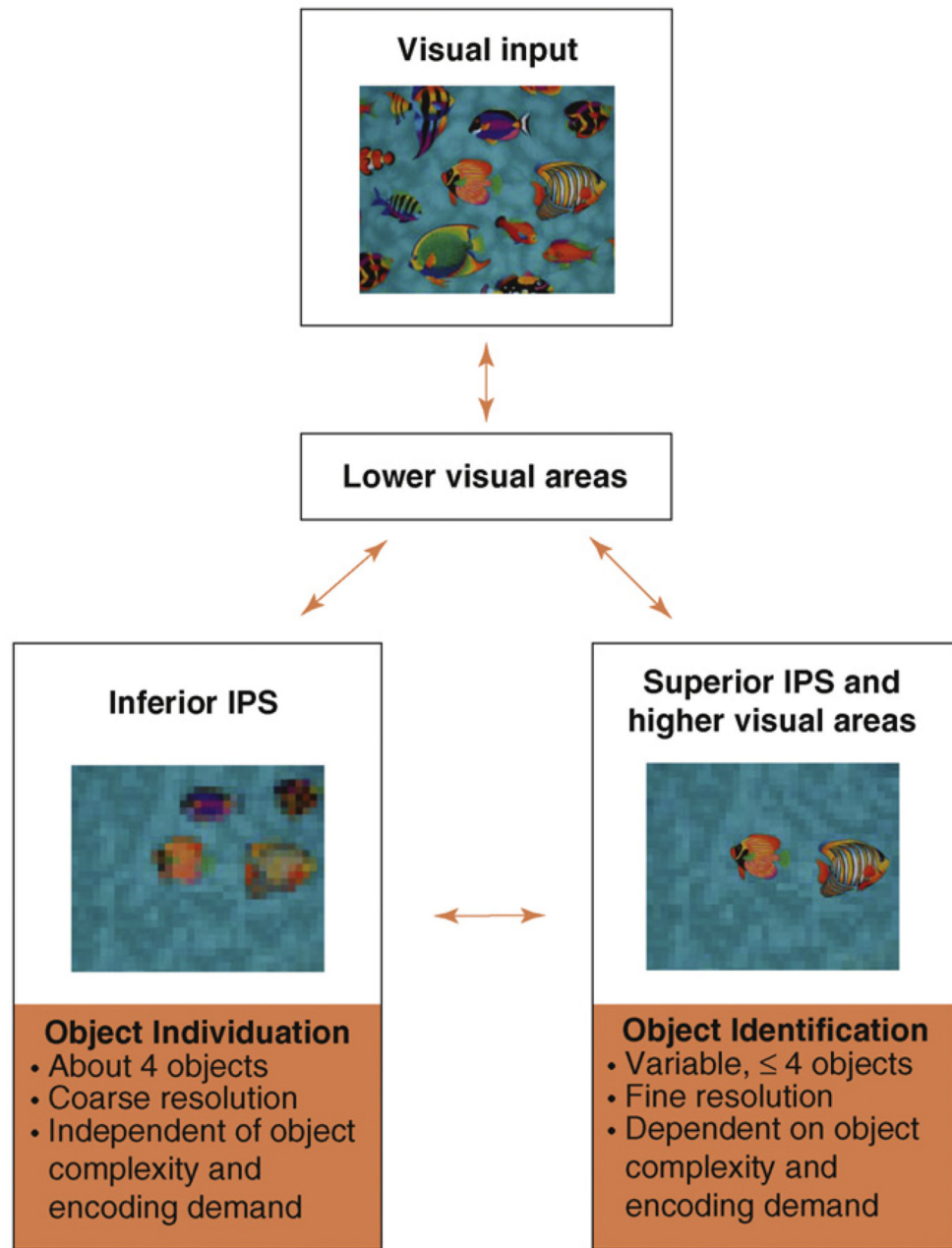
References

1. Kahneman D, et al. The reviewing of object files: object-specific integration of information. *Cognit. Psychol.* 1992; 24:175–219. [PubMed: 1582172]
2. Xu Y, Chun MM. Dissociable neural mechanism supporting visual short-term memory for objects. *Nature.* 2006; 440:91–95. [PubMed: 16382240]
3. Xu Y, Chun MM. Visual grouping in human parietal cortex. *Proc. Natl. Acad. Sci. U. S. A.* 2007; 104:18766–18771. [PubMed: 17998539]
4. Xu Y. The role of the superior intra-parietal sulcus in supporting visual short-term memory for multi-feature objects. *J. Neurosci.* 2007; 27:11676–11686. [PubMed: 17959810]
5. Xu Y. Representing connected and disconnected shapes in human inferior intra-parietal sulcus. *Neuroimage.* 2008; 40:1849–1856. [PubMed: 18353688]
6. Xu Y. Distinctive neural mechanisms supporting visual object individuation and identification. *J. Cogn. Neurosci.* 2009; 21:511–518. [PubMed: 18510449]
7. Pylyshyn Z. Some primitive mechanisms of spatial attention. *Cognition.* 1994; 50:363–384. [PubMed: 8039369]

8. Rensink RA. The dynamic representation of scenes. *Vis. Cogn.* 2000; 7:17–42.
9. Cowan N. The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behav. Brain Sci.* 2001; 24:87–114. [PubMed: 11515286]
10. Pylyshyn ZW, Storm RW. Tracking multiple independent targets: evidence for a parallel tracking mechanism. *Spat. Vis.* 1988; 3:179–197. [PubMed: 3153671]
11. Cavanagh P, Alvarez GA. Tracking multiple targets with multifocal attention. *Trends Cogn. Sci.* 2005; 9:349–354. [PubMed: 15953754]
12. Alvarez GA, Cavanagh P. The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychol. Sci.* 2004; 15:106–111. [PubMed: 14738517]
13. Todd JJ, Marois R. Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature.* 2004; 428:751–754. [PubMed: 15085133]
14. Itti L, Koch C. A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Res.* 2000; 40:1489–1506. [PubMed: 10788654]
15. Corbetta M, Shulman GL. Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 2002; 3:215–229.
16. Ungerleider, LG.; Mishkin, M. Two cortical visual systems. In: Ingle, DJ., et al., editors. *Analysis of Visual Behavior.* MIT Press; 1982. p. 549–586.
17. Sereno AB, Maunsell JHR. Shape selectivity in primate lateral intraparietal cortex. *Nature.* 1998; 395:500–503. [PubMed: 9774105]
18. Konen CS, Kastner S. Two hierarchically organized neural systems for object information in human visual cortex. *Nat. Neurosci.* 2008; 11:224–231. [PubMed: 18193041]
19. Merigan WH, Maunsell JH. How parallel are the primate visual pathways? *Annu. Rev. Neurosci.* 1993; 16:369–402. [PubMed: 8460898]
20. Awh E, et al. Visual working memory represents a fixed number of items, regardless of complexity. *Psychol. Sci.* 2007; 18:622–628. [PubMed: 17614871]
21. Rauschenberger R, Yantis S. Masking unveils pre-amodal completion representation in visual search. *Nature.* 2001; 410:369–372. [PubMed: 11268213]
22. Zemel RS, et al. Experience-dependent perceptual grouping and object-based attention. *J. Exp. Psychol. Hum. Percept. Perform.* 2002; 28:202–217.
23. Vecera SP, O’Reilly RC. Figure-ground organization and object recognition processes: an interactive account. *J. Exp. Psychol. Hum. Percept. Perform.* 1998; 24:441–462. [PubMed: 9554093]
24. Sagi D, Julesz B. “What” and “where” in vision. *Science.* 1985; 228:1217–1219. [PubMed: 4001937]
25. Leslie AM, et al. Indexing and the object concept: developing “what” and ‘where’ systems. *Trends Cogn. Sci.* 1998; 2:10–18. [PubMed: 21244957]
26. Luck SJ, Vogel EK. The capacity of visual working memory for features and conjunctions. *Nature.* 1997; 390:279–281. [PubMed: 9384378]
27. Zhang W, Luck SJ. Discrete fixed-resolution representations in visual working memory. *Nature.* 2008; 453:233–235. [PubMed: 18385672]
28. Wheeler M, Treisman AM. Binding in short-term visual memory. *J. Exp. Psychol. Gen.* 2002; 131:48–64. [PubMed: 11900102]
29. Xu Y. Integrating color and shape in visual short-term memory for objects with parts. *Percept. Psychophys.* 2002; 64:1260–1280. [PubMed: 12519024]
30. Xu Y. Limitations in object-based feature encoding in visual short-term memory. *J. Exp. Psychol. Hum. Percept. Perform.* 2002; 28:458–468. [PubMed: 11999866]
31. Xu Y. Encoding objects in visual short-term memory: the roles of feature proximity and connectedness. *Percept. Psychophys.* 2006; 68:815–828. [PubMed: 17076349]
32. Bays PM, Husain M. Dynamic shifts of limited working memory resources in human vision. *Science.* 2008; 321:851–854. [PubMed: 18687968]
33. Vogel EK, Machizawa MG. Neural activity predicts individual differences in visual working memory capacity. *Nature.* 2004; 428:748–751. [PubMed: 15085132]

34. Jonides J, et al. Processes of working memory in mind and brain. *Curr. Dir. Psychol. Sci.* 2005; 14:2–5.
35. Pasternak T, Greenlee MW. Working memory in primate sensory systems. *Nat. Rev. Neurosci.* 2005; 6:97–107. [PubMed: 15654324]
36. Duncan J. Selective attention and the organization of visual information. *J. Exp. Psychol. Gen.* 1984; 113:501–517. [PubMed: 6240521]
37. Egly R, et al. Shifting visual attention between objects and locations: evidence for normal and parietal lesion subjects. *J. Exp. Psychol. Gen.* 1994; 123:161–177. [PubMed: 8014611]
38. Behrmann M, et al. Object-based attention and occlusion: evidence from normal participants and a computational model. *J. Exp. Psychol. Hum. Percept. Perform.* 1998; 24:1011–1036. [PubMed: 9706708]
39. Scholl BJ. Objects and attention: the state of the art. *Cognition.* 2001; 80:1–46. [PubMed: 11245838]
40. Mattingley JB, et al. Preattentive filling-in of visual surfaces in parietal extinction. *Science.* 1997; 275:671–674. [PubMed: 9005854]
41. Coslett HB, Saffran E. Simultanagnosia: to see but not two see. *Brain.* 1991; 114:1523–1545. [PubMed: 1884165]
42. Friedman-Hill SR, et al. Parietal contributions to visual feature binding: evidence from a patient with bilateral lesions. *Science.* 1995; 269:853–855. [PubMed: 7638604]
43. Baharmi B. Object property encoding and change blindness in multiple object tracking. *Vis. Cogn.* 2003; 10:949–963.
44. Horowitz TD, et al. Tracking unique objects. *Percept. Psychophys.* 2007; 69:172–184. [PubMed: 17557588]
45. Culham JC, et al. Cortical fMRI activation produced by attentive tracking of moving targets. *J. Neurophysiol.* 1998; 80:2657–2670. [PubMed: 9819271]
46. Kaufman EL, et al. The discrimination of visual number. *Am. J. Psychol.* 1949; 62:498–525. [PubMed: 15392567]
47. Trick LM, Pylyshyn ZW. Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychol. Rev.* 1994; 101:80–102. [PubMed: 8121961]
48. Piazza M, et al. Single-trial classification of parallel pre-attentive and serial attentive processes using functional magnetic resonance imaging. *Proc. Biol. Sci.* 2003; 270:1237–1245. [PubMed: 12816636]
49. Pylyshyn ZW. The role of location indexes in spatial perception: a sketch of the FINST spatial-index model. *Cognition.* 1989; 32:65–97. [PubMed: 2752706]
50. Alvarez GA, Franconeri SL. How many objects can you attentively track? Evidence for a resource-limited tracking mechanism. *J. Vision.* 2007; 7 DOI:10.1167/7.13.14.
51. Sereno MI, et al. Mapping of contralateral space in retinotopic coordinates by a parietal cortical area in humans. *Science.* 2001; 294:1350–1354. [PubMed: 11701930]
52. Silver MA, et al. Topographic maps of visual spatial attention in human parietal cortex. *J. Neurophysiol.* 2005; 94:1358–1371. [PubMed: 15817643]
53. Swisher JD, et al. Visual topography of human intraparietal sulcus. *J. Neurosci.* 2007; 27:5326–5337. [PubMed: 17507555]
54. Nieder A. Counting on neurons: the neurobiology of numerical competence. *Nat. Rev. Neurosci.* 2005; 6:177–190. [PubMed: 15711599]
55. Sugrue LP, et al. Choosing the greater of two goods: neural currencies for valuation and decision making. *Nat. Rev. Neurosci.* 2005; 6:1–13.
56. Glimcher PW. The neurobiology of visual saccadic decision making. *Annu. Rev. Neurosci.* 2003; 26:133–179. [PubMed: 14527268]
57. Wagner AD, et al. Parietal lobe contributions to episodic memory retrieval. *Trends Cogn. Sci.* 2005; 9:445–453. [PubMed: 16054861]
58. Macaluso E, Driver J. Multisensory spatial interactions: a window on functional integration in the human brain. *Trends Neurosci.* 2005; 28:264–271. [PubMed: 15866201]

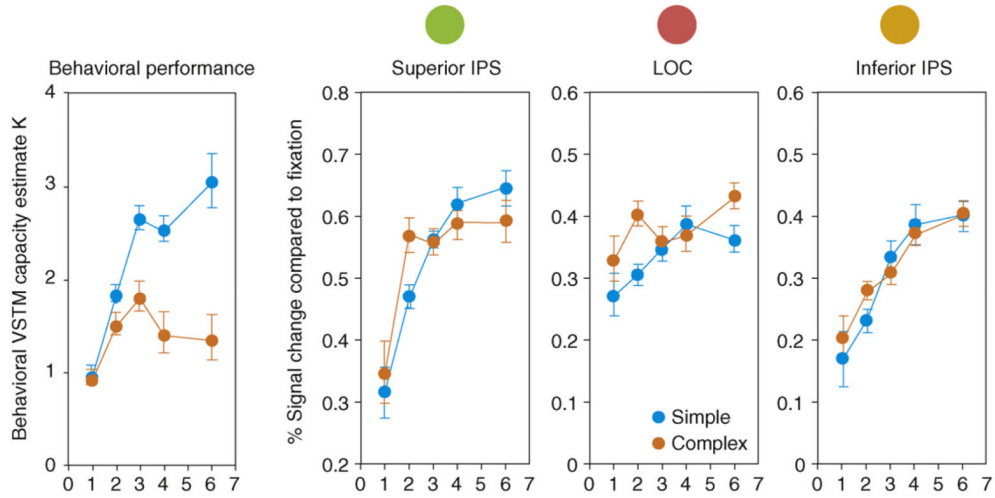
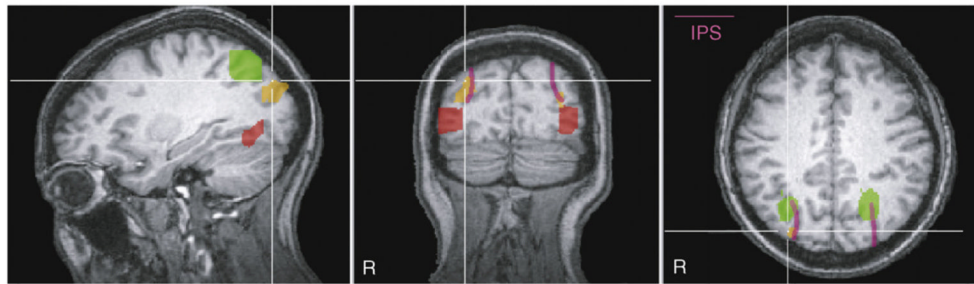
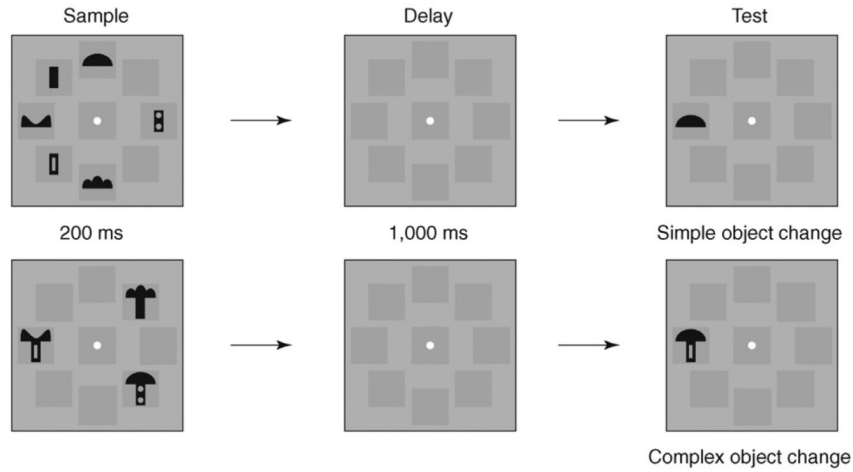
59. Gottlieb J. From thought to action: the parietal cortex as a bridge between perception, action, and cognition. *Neuron*. 2007; 53:9–16. [PubMed: 17196526]
60. Serences, J., et al. Parietal mechanisms of switching and maintaining attention to locations, objects, and features. In: Itti, L., et al., editors. *Neurobiology of Attention*. Academic Press; 2005. p. 35–41.
61. Epstein R, Kanwisher N. A cortical representation of the local visual environment. *Nature*. 1998; 392:598–601. [PubMed: 9560155]
62. Xu F, Carey S. Infants' metaphysics: the case of numerical identity. *Cognit. Psychol.* 1996; 30:111–153. [PubMed: 8635312]
63. Káldy Z, Leslie A. A memory span of one? Object identification in 6.5-month-old infants. *Cognition*. 2005; 57:153–177.
64. Green CS, Bavelier D. Action video games modify visual attention. *Nature*. 2003; 423:534–537. [PubMed: 12774121]
65. Green CS, Bavelier D. Enumeration versus object tracking: Insights from video game players. *Cognition*. 2006; 101:217–245. [PubMed: 16359652]
66. Scolaro M, et al. Perceptual expertise enhances the resolution but not the number of representations in working memory. *Psychon. Bull. Rev.* 2008; 15:215–222. [PubMed: 18605506]
67. Curby K, et al. A visual short-term memory advantage for objects of expertise. *J. Exp. Psychol. Hum. Percept. Perform.* 2009; 35:94–107. [PubMed: 19170473]
68. Gauthier I, et al. Expertise for cars and birds recruits brain areas involved in face recognition. *Nat. Neurosci.* 2000; 3:191–197. [PubMed: 10649576]
69. Op de Beeck H, et al. Discrimination training alters object representations in human extrastriate cortex. *J. Neurosci.* 2006; 26:13025–13036. [PubMed: 17167092]
70. Jiang X, et al. Categorization training results in shape- and category-selective human neural plasticity. *Neuron*. 2007; 53:891–903. [PubMed: 17359923]



TRENDS in Cognitive Sciences

Figure 1.

Our neural object-file theory posits that there is an object individuation stage in which a fixed number of about four objects are first individuated and selected from a crowded scene based on their spatial information by the inferior intra-parietal sulcus (IPS), and an object identification stage in which a subset of the selected objects are encoded into greater detail by higher visual areas and the superior IPS. Although object individuation normally precedes object identification, results of object identification can refine or modify what is selected during initial object individuation, such as in perceiving occluded objects.

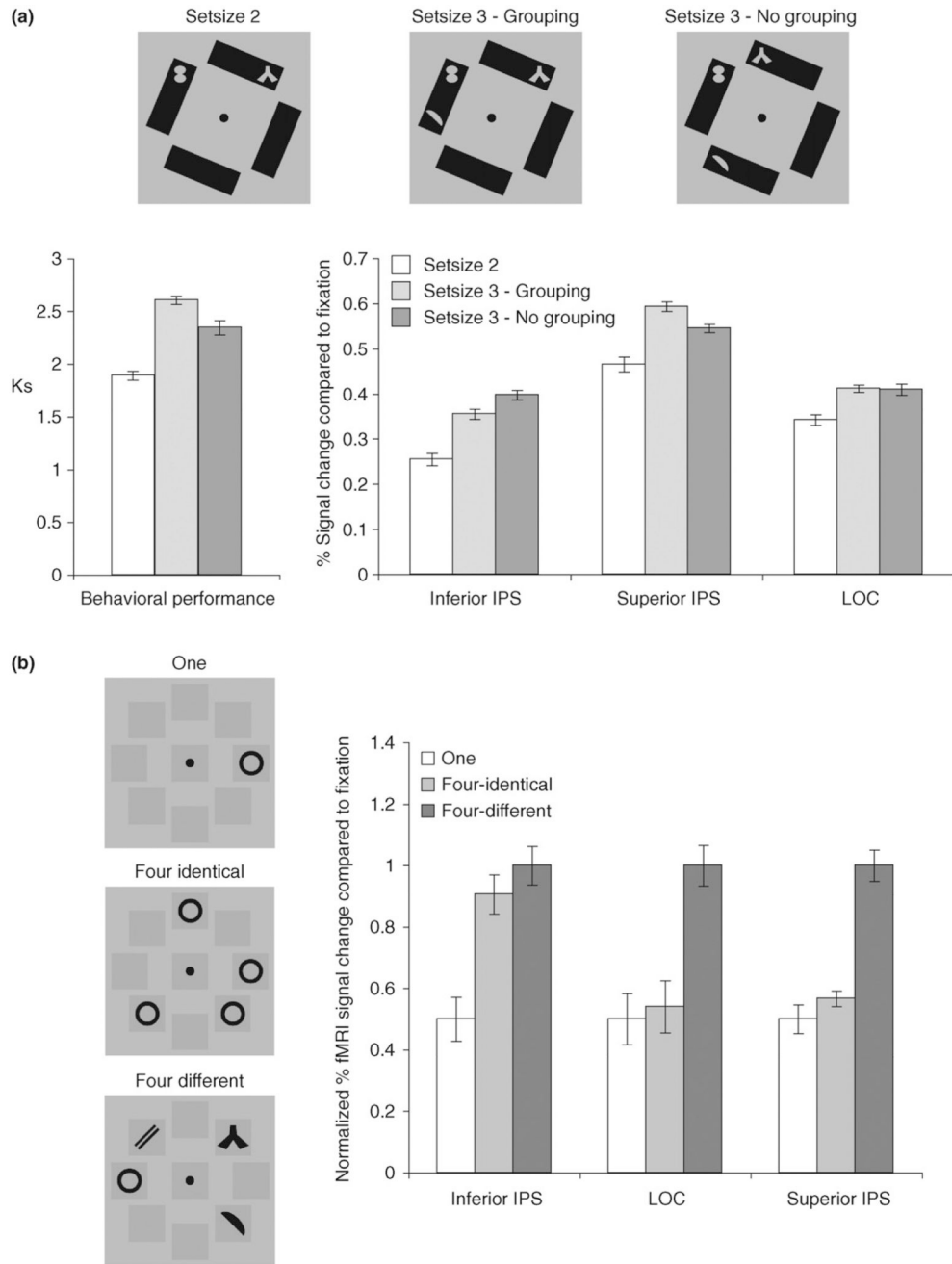


TRENDS in Cognitive Sciences

Figure 2.

In this experiment [5], observers viewed a variable number of briefly appearing black shapes and performed change detection after a brief delay. The black shapes were either simple or complex, constructed by attaching two simple shapes together. Superior and inferior IPS and LOC regions of interest (ROIs) are shown here. Although inferior IPS response increased with display set-size and plateaued at about set size 4 regardless of object complexity, superior IPS and LOC responses tracked VSTM capacity as determined by object complexity. Thus, corresponding to object individuation, the inferior IPS selects a fixed number of approximately four objects via their spatial locations, and corresponding to object

identification, the superior IPS and the LOC encode the shape features of a subset of the selected objects into great detail.



TRENDS in Cognitive Sciences

Figure 3.

(a) In this experiment [6], observers viewed 2 or 3 briefly appearing grey shapes and performed change detection. The display background determined shape grouping although it was task irrelevant. Inferior IPS response was lower for the three grouped than for the three ungrouped shapes. This grouping effect reversed direction in behavioral VSTM capacity estimates and superior IPS responses (but was absent in the LOC), showing a VSTM grouping benefit. **(b)** In this experiment [4], observers viewed either one black shape, four identical black shapes or four different black shapes in a sample display and performed change detection. Although responses for the two four-object conditions were similar in the inferior IPS, those for the one and the four-identical conditions were similar in the LOC and

the superior IPS. These results support object individuation in the inferior IPS and object identification in the LOC and the superior IPS (normalized functional magnetic resonance imaging [fMRI] results were shown here, in which the averaged responses for the one and the four-different conditions were anchored to be 0.5 and 1, respectively).