ELSEVIER

Learning and neural plasticity in visual object recognition

Zoe Kourtzi[1] and James J DiCarlo[2]

The capability of the adult primate visual system for rapid and accurate recognition of targets in cluttered, natural scenes far surpasses the abilities of state-of-the-art artificial vision systems. Understanding this capability remains a fundamental challenge in visual neuroscience. Recent experimental evidence suggests that adaptive coding strategies facilitated by underlying neural plasticity enable the adult brain to learn from visual experience and shape its ability to integrate and recognize coherent visual objects.

Addresses
[1] School of Psychology, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK
[2] McGovern Institute for Brain Research and Department of Brain and Cognitive Sciences, MIT, Cambridge, MA, 02139, USA

Corresponding author: Kourtzi, Zoe (z.kourtzi@bham.ac.uk)

Introduction

Detecting and recognizing meaningful objects in complex environments is a crucial skill that underlies a range of behaviours, from identifying predators and prey and recognizing edible and poisonous foods, to diagnosing tumours on medical images and finding familiar faces in a crowd. In humans and other visual primates, these processes operate quickly, automatically and effortlessly, and are thus easily taken for granted. However, the computational challenges of visual recognition are far from trivial. In particular, the recognition of coherent meaningful objects entails integration at different levels of visual complexity, from local contours to complex objects, and representations that are highly tolerant of identity-preserving image changes (e.g. changes in position, size, pose or background clutter).

A wide range of methods provide converging evidence that neuronal processes supporting object recognition are coarsely localized in the ventral visual stream [1], which has a rough hierarchy of cortical processing stages (V1 [primary visual cortex] → V2 → V4 → PIT [posterior inferior temporal cortex] → AIT [anterior inferior temporal cortex]). The highest stages of this stream (i.e. anterior inferior temporal cortex [AIT] in the monkey, and lateral occipital complex [LOC] in the human) are thought to convey neuronal signals that are well suited to support object recognition directly. In particular, unlike earlier visual areas, patterns of neuronal activity in these regions explicitly convey object identity, in that object identity can be directly extracted from those populations, even in the face of identity-preserving image changes [2–5]. But how are such useful neuronal representations constructed in the brain? How do neuronal connections become wired up and modified so that neurons respond to complex combinations of simple image features and are sensitive to subtle changes in object identity, yet are relatively insensitive to large, identity-preserving image changes? How does the brain even know what an 'object' is in the first place? Do these neuronal representations code for all possible objects, or do they just represent objects that are behaviourally relevant or often encountered in the environment? At the core of these issues are fundamental questions about the role of visual experience and learning in the establishment and maintenance of the neuronal representations that support complex object recognition (see [6] for an earlier review).

At a theoretical level, there is growing appreciation of the potentially powerful role of learning in establishing robust representations crucial for object recognition [7–12]. Experimentally, the role of learning in object representation can be approached by studying developing visual systems [13], visual systems that have been deprived of experience during early life [14,15] and the role of visual experience in adults [16–18]. All of these approaches have been used to gain new insight into the role of learning in feature and object representation. For the purposes of this review, we focus on experience-based plasticity in the adult visual system. In particular, many psychophysical studies in adults have shown learning-dependent changes in discrimination and recognition using stimuli ranging from simple features, such as oriented lines and gratings [19], to complex objects [20]. Recent neurophysiological [21–25,26•,27–30,31••,32–34] and functional magnetic resonance imaging (fMRI) [35–39,40•,41•] studies have focused on elucidating the loci of brain plasticity and changes in neuronal responses that underlie this visual learning. Here, we briefly review these advances and propose that experience-based plasticity across multiple stages of visual analysis bolsters selective, robust representations of visual objects, and thus directly underlies the perceptual integration of local features into coherent meaningful objects and their recognition in the complex environments we inhabit.

Learning to put an object together

The ability to build neuronal representations that are highly selective for visual pixel combinations ('features') and feature combinations ('objects') in addition to being highly tolerant to identity-preserving image changes is the computational crux of object recognition [42–44], and the hallmark of primate vision. At the most basic level, object recognition requires the visual system to discriminate among different patterns of visual input (e.g. the letter 'A' among all other possible letters and objects). As this discrimination cannot be solved simply by monitoring the output of any one photoreceptor (or even one lateral geniculate nucleus [LGN] or V1 neuron), a solution to this problem requires some 'binding' of responses from neurons at early visual stages. Standard computational approaches propose that 'binding' is achieved by synaptically combining inputs from neurons at early visual stages (e.g. through a thresholded weighted sum of such inputs). As a result, neurons at higher stages of the visual system are tuned to patterns of increasing complexity until the required pattern discrimination can be supported [45]. Because pattern discrimination must be performed in the real world, this selectivity for combinations of visual features should have robust tolerance for image changes that produce profound transformations in the visual input without modifying object identity (e.g. position, size, pose or background clutter) [42,44]. Computational models [11,43,46] have shown that such explicit object representations can be built using neuronal connections that group together similar features regardless of image changes. In this view, the process of constructing neurons that are both tuned for complex configurations of simpler visual features (e.g. pixels, edges) and relatively insensitive to some types of image changes is equivalent to the process of defining what an 'object' is in the first place. The experimental challenge is to understand how this is done in the brain.

Learning selectivity

How does the brain figure out how to appropriately connect neurons together and weight their inputs so that selective neuronal representations are ultimately obtained? Plasticity of neuronal connections along with appropriate learning 'rules' is one obvious potential mechanism. For example, learning to respond selectively to an object might amount to learning which simpler image features tend to often co-occur in the world [47,48] (Figure 1a). At a mechanistic level, this could result if inputs conveying simpler image features are brought together at downstream neurons that respond non-linearly to those inputs (e.g. respond only if feature A and feature B are both present). An iteration of this strategy at each level of the visual system would result in progressively more complex preferred stimulus configurations. Competitive mechanisms could ensure efficient coverage of the subspace of possible images that is spanned by natural images [49]. Recent neurophysiological studies provide evidence for such learning. In particular, neurons in monkey IT cortex show enhanced selectivity after training for novel objects [23,25,32], holistic multiple-part configurations [29] and even physically unrelated pairs of shapes [24,50•]. The time-course of changes in some of these neurons parallel that of learning [51], suggesting a strong link between underlying neuronal plasticity and behavioural improvement. Furthermore, learning can shape the assignment of novel objects into classes [52] by modulating the selectivity of neurons in the inferior temporal and frontal cortex for

Figure 1
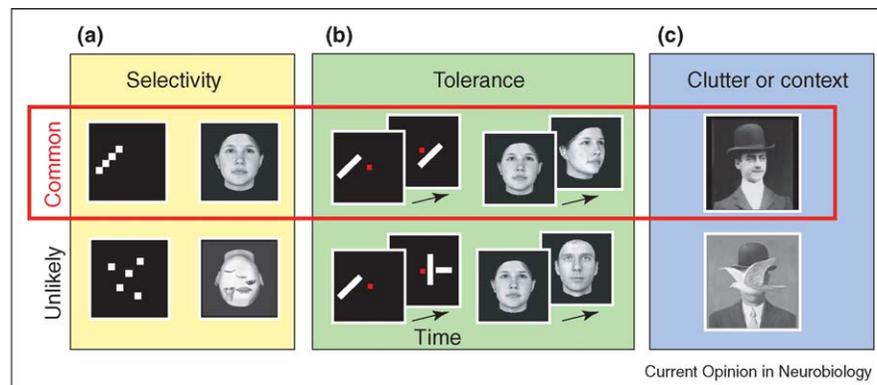
Key neuronal response properties required for robust object recognition might ultimately be acquired and maintained through visual experience. **(a)** Image arrangements that are well represented by the visual system are those that are commonly encountered in the world (compare the upper with the lower images in the left panel). **(b)** Similarly, image transformations that the visual system is able to tolerate robustly are those that are often experienced across short time intervals. For example, the upper image sequences show translation and pose change respectively without changes in object identity. The lower image sequences show changes in object identity across short time intervals, a situation that is rarely encountered, but can alter the tolerance of object recognition [54,58•]. **(c)** Items in a scene that are not part of an object can make object recognition more difficult (clutter), but if those items are often seen with the object (e.g. hat, shirt and jacket are often seen with a face), the items themselves might be sufficient for object representation and perception [68].

features crucial for these categorization processes [27,30,53[•]].

Learning tolerance

To date, most studies of visual learning have focused on changes in neuronal or behavioural selectivity. However, as described above, simply learning selectivity is not enough to create useful object representations — that is, selective object representations must be tolerant to image changes (e.g. object position, size and pose). But how does the visual system know which neurons to connect (or, equivalently, how to weight those connections) to enable this tolerance? Again, learning from the statistical regularities of the natural world has been proposed as a potential solution [8–11,54]. One central idea is that features and objects in the world do not tend to jump in and out of existence, but they have temporal continuity, in that an object seen at one instant in time will probably be seen in the next instant, but perhaps in a different position, size or pose (Figure 1b). Another related idea is that once semi-tolerant representations are established, the later appearance of a feature or object that is similar to that representation, but differs slightly in (e.g.) position, scale, or view, can re-activate the initial representation and enable tolerance learning without temporal continuity [55]. Although 'tolerance learning' is of great computational importance, and there is some behavioural [19] and circumstantial neuronal evidence [32,53[•],56] suggesting that tolerance is not automatic, its neural basis remains largely unknown.

Recently, however, psychophysical studies have directly demonstrated that targeted disruptions of the temporal continuity of an object result in disruptions in object perception consistent with tolerance learning [54,57,58[•]]. For example, Cox *et al.* [58[•]] recently found that even the most fundamental type of recognition tolerance — the ability to recognize an object despite its position on the retina — can be predictably modified by visual experience. In particular, changes to object identity during normal eye movements that bring the 'object' from one retinal position to another over short time intervals disrupt later recognition of the object across those same retinal positions. But there is also evidence that view tolerance can be learned without temporal continuity [59]. An important goal of ongoing and future work is to elucidate neuronal changes in the ventral stream in the context of tolerance learning.

Learning objects in clutter

Beyond learning selectivity and tolerance for object identity, the visual system must learn to detect objects in the real world, in which they are seen in clutter and context [60,61]. During a course of training, observers can learn distinctive target features by using information (image regularities) crucial for target detection more efficiently and by suppressing background noise [62–65,66[•]]. In particular, learning has been suggested to enhance correlations among neurons responding to the features of target patterns while de-correlating neural responses to target and background patterns. As a result, input (stimulus) redundancy is reduced and target salience is enhanced [33], supporting the efficient detection and identification of objects in cluttered scenes [67].

It should be noted however, that the background 'stuff' in which an object is embedded in a scene should not always be viewed as 'clutter' that must be ignored. Indeed, that 'clutter' can include features that, although not perfectly correlated with the object, are more often than not seen with the object (i.e. context), and thus might aid the detection and recognition of the object (Figure 1c). Thus, the visual system might learn to incorporate these clues in its object representations [68,69]. In addition, the 'clutter' will typically include other objects that might also need to be detected and recognized. Understanding how the visual system represents multiple objects simultaneously [70,71] is crucial for unravelling the mechanisms that mediate successful interactions in complex, dynamic environments.

Neuronal plasticity underlying visual object learning

Studies demonstrating experience-dependent changes in the selectivity and tolerance of high-level neuronal representations (e.g. IT) beg the question of the locus and nature of these changes, as improvement at higher stages of visual analysis might be inherited from changes at one or more earlier stages.

It has often been suggested [19,72,73] that the key plasticity locus in simple feature 'perceptual learning' is likely to be in early visual stages, as this learning is somewhat confined to the trained retinal position. That is, changes in the receptive field tuning properties of neurons in V1 might account for the specificity of learning effects to the trained visual field position and trained stimulus attribute. Indeed, recent imaging studies [37,40[•],74] provide evidence for the involvement of V1 in object feature learning. However, neurophysiological evidence for the contribution of V1 in behavioural improvement after training on visual discrimination remains controversial [21,22]. There is some evidence for sharpening of orientation tuning after training [22], but no evidence for changes in the size of the cortical representation or the receptive field properties of neurons in V1 [21,75]. One possibility is that V1 learning effects can be detected in the average response of large numbers of neurons, as measured by fMRI, but are very small at the level of the individual neurons. Another possibility is that they reflect task-dependent changes in intra- and inter-area connectivity.

Recently, two studies combining psychophysics and fMRI [41[•],76[•]] examined the relationship between shape

learning and experience-dependent reorganization across stages of visual processing. The results of Sigman et al. [76•] suggest that shape representation might shift from higher to early visual areas, which support rapid and automatic search and detection in visual cluttered scenes independently of attentional control. These findings are consistent with the suggestion that learning moulds object representations not only by enhancing the processing of feature detectors with increasing complexity in a bottom-up manner but also in a top-down manner taking into account the relevant task dimensions and demands. In particular, it has been suggested that learning begins at higher visual areas for easy tasks and proceeds to early retinotopic areas that have higher resolution for finer and more difficult discriminations [77••]. Kourtzi et al. [41•] provide evidence that these distributed plasticity mechanisms are adaptable to natural image regularities that determine the salience of targets in cluttered scenes. In particular, their results suggest that opportunistic learning [63] of salient targets in natural scenes is mediated by sparser feature coding at higher stages of visual analysis, whereas learning of camouflaged targets is implemented by mechanisms that enhance the segmentation and recognition of ambiguous targets in both early and higher visual areas.

One of the main advantages of fMRI is that it provides global brain coverage and, thus, it is a highly suitable method for studying learning-dependent changes across stages of analysis in the visual system. However, experience-dependent activation changes in fMRI studies could be the result of changes in the numbers or the gain of neurons recruited for processing of a stimulus in the context of a task. As imaging studies measure activation at the large scale of neural populations rather than at the scale of single neurons, they cannot discern these different neural plasticity mechanisms. Recent neurophysiological studies [26•,28] have shed light on to cortical reorganization mechanisms at the level of the single neuron when monkeys learn to discriminate images of natural scenes presented in noise. These studies show that learning enhances the selective processing of crucial features for the detection of object targets in early occipito-temporal areas. By contrast, learning appears to facilitate efficient object processing independent of background noise in the prefrontal cortex. These findings suggest that learning in different cortical areas bolsters functions that are important for different tasks, ranging from the bottom-up detection and integration of target features in cluttered scenes across visual occipitotemporal areas to the top-down selection of familiar objects in the prefrontal cortex. Consistent with top-down approaches to visual processing, recent neuroimaging studies suggest that learning might enhance the functional interactions between occipitotemporal areas that encode physical stimulus experiences and parieto-frontal circuits that represent our perceptual interpretations of the world

[38,78,79]. Future studies combining fMRI and simultaneous chronic recordings from these areas will provide novel insights for understanding both bottom-up and top-down mechanisms for experience-dependent reorganization at the level of inter- and intra- area networks.

In summary, the current experimental evidence suggests that there is no single locus of brain plasticity underlying visual learning. These findings are consistent with computational approaches proposing that associations between features that mediate the recognition of familiar objects might occur across stages of visual analysis, from orientation detectors in the primary visual cortex to occipitotemporal neurons tuned to object parts and views [7,8,43]. At the neuronal level, learning could be implemented by changes in core feedforward neuronal processing, especially at higher visual stages [31••], or by changes in the interactions between object analysis centres in temporal and frontal cortical areas and local connections in the primary visual cortex based on top-down feedback mechanisms. For example, learning has been suggested to modulate neuronal sensitivity in the early visual areas by modulating networks of lateral interactions and through feedback connections from higher visual areas [17,74,75,80–82]. Such changes in the connectivity of visual analysis circuits might be adaptive and efficient compared with changes in core feedforward visual processing (e.g. receptive fields) that might have deleterious consequences for the visual processing of the trained stimuli in another context or task. Current research directions focus on further understanding the effects of stimulus and task demands on learning across stages of visual analysis, the relative time courses of learning-dependent changes, and the underlying neuronal responses and network interactions that change to enable learning to occur while not disrupting general visual processing.

Conclusions

Visual object perception and recognition in cluttered, natural scenes poses a series of computational challenges to the adult visual system, from the detection of image regularities to binding contours, parts and features into coherent objects, recognizing them independent of image changes (e.g. position, scale, pose, clutter) and assigning them to abstract categories. This review highlights the potentially fundamental role of learning in solving some of these challenges. What general conclusions are we to take from the experimental evidence available so far? First, the adult visual system is clearly plastic, in terms of both behavioural improvements and changes in neuronal responses. Second, there is no single locus of plasticity in the visual system that is the exclusive site underlying object learning. On the contrary, in most cases learning modifies visual representations for features and objects by modulating processing across multiple cortical levels. Third, learning does not always result in simple, static

changes to core feedforward visual processing — instead, changes can be dynamic and task dependent. Thus, understanding object learning cannot be divorced from the context of the computational problems faced by the visual system in complex environments; that is, learning robust object representations depends on the stimulus conditions and the task demands. An important goal for future work is to understand the effects of these factors on optimal computations and the neuronal correlates of visual learning. Finally, the relationship between the neural mechanisms that mediate adult, experience-dependent plasticity and developmental plasticity is intriguing and remains largely unknown. Although the adult visual system is remarkably powerful at representing and distinguishing among objects even the very first time they are seen, this cannot rule out a potentially crucial role of visual experience in the establishment of such representations. On the contrary, it is surprising that small amounts of adult visual experience, as reviewed here, produce measurable changes in both behaviour and neuronal representations, even when superimposed on a lifetime of natural experience. To the extent that adult visual learning shares computational and perhaps even some mechanistic commonalities with the developing visual system, understanding experience-based plasticity could reveal the key principles that underlie our remarkable ability for robust object recognition.

Acknowledgements

References and recommended reading

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

- • of special interest
- •• of outstanding interest

1. Ungerleider LG, Mishkin M: **Two cortical visual systems**. In *Analysis of Visual Behavior.* Edited by Ingle DJ, Goodale MA, Mansfield RJW. M.I.T. Press; 1982:549-585.

2. Rolls ET: **Functions of the primate temporal lobe cortical visual areas in invariant visual object and face recognition**. *Neuron* 2000, **27**:205-218.

3. Hung CP, Kreiman G, Poggio T, DiCarlo JJ: **Fast readout of object identity from macaque inferior temporal cortex**. *Science* 2005, **310**:863-866.

4. Quiroga RQ, Reddy L, Kreiman G, Koch C, Fried I: **Invariant visual representation by single neurons in the human brain**. *Nature* 2005, **435**:1102-1107.

5. Grill-Spector K, Malach R: **The human visual cortex**. *Annu Rev Neurosci* 2004, **27**:649-677.

6. Wallis G, Bulthoff H: **Learning to recognize objects**. *Trends Cogn Sci* 1999, **3**:22-31.

7. Poggio T: **A theory of how the brain might work**. *Cold Spring Harb Symp Quant Biol* 1990, **55**:899-910.

8. Wallis G, Rolls ET: **Invariant face and object recognition in the visual system**. *Prog Neurobiol* 1997, **51**:167-194.

9. Foldiak P: **Learning invariance from transformation sequences**. *Neural Comput* 1991, **3**:194-200.

10. Wiskott L, Sejnowski TJ: **Slow feature analysis: unsupervised learning of invariances**. *Neural Comput* 2002, **14**:715-770.

11. Ullman S, Soloviev S: **Computation of pattern invariance in brain-like structures**. *Neural Netw* 1999, **12**:1021-1036.

12. Karklin Y, Lewicki MS: **A hierarchical Bayesian model for learning nonlinear statistical regularities in nonstationary natural signals**. *Neural Comput* 2005, **17**:397-423.

13. Rodman HR: **Development of inferior temporal cortex in the monkey**. *Cereb Cortex* 1994, **4**:484-498.

14. Blakemore C, Van Sluyters RC: **Innate and environmental factors in the development of the kitten's visual cortex**. *J Physiol* 1975, **248**:663-716.

15. Sinha P, Bouvrie JVaS: **Object concept learning: observations in congenitally blind children and a computational model**. *Neurocomputing* 2006. in press.

16. Goldstone RL: **Perceptual learning**. *Annu Rev Psychol* 1998, **49**:585-612.

17. Gilbert CD, Sigman M, Crist RE: **The neural basis of perceptual learning**. *Neuron* 2001, **31**:681-697.

18. Schyns PG, Goldstone RL, Thibaut JP: **The development of features in object concepts**. *Behav Brain Sci* 1998, **21**:1-17, discussion 17-54.

19. Fahle M: **Perceptual learning: a case for early selection**. *J Vis* 2004, **4**:879-890.

20. Fine I, Jacobs RA: **Comparing perceptual learning tasks: a review**. *J Vis* 2002, **2**:190-203.

21. Ghose GM, Yang T, Maunsell JH: **Physiological correlates of perceptual learning in monkey V1 and V2**. *J Neurophysiol* 2002, **87**:1867-1888.

22. Schoups A, Vogels R, Qian N, Orban G: **Practising orientation identification improves orientation coding in V1 neurons**. *Nature* 2001, **412**:549-553.

23. Kobatake E, Wang G, Tanaka K: **Effects of shape-discrimination training on the selectivity of inferotemporal cells in adult monkeys**. *J Neurophysiol* 1998, **80**:324-330.

24. Miyashita Y: **Cognitive memory: cellular and network machineries and their top-down control**. *Science* 2004, **306**:435-440.

25. Rolls ET: **Learning mechanisms in the temporal lobe visual cortex**. *Behav Brain Res* 1995, **66**:177-185.

26. Rainer G, Lee H, Logothetis NK: **The effect of learning on the**
• **function of monkey extrastriate visual cortex**. *PLoS Biol* 2004, **2**:E44.
This study, in combination with previous findings (Rainer *et al.*, [28]), suggests that different neural plasticity mechanisms in early visual areas (V4) and the prefrontal cortex mediate enhanced behavioural performance in discriminating degraded images after training. In particular, stimulus degradation by increasing noise resulted in decreased responses of V4 neurons to novel images, whereas learning enhanced their responses to degraded stimuli. By contrast, the number of prefrontal neurons that responded to undegraded images decreased as these images became familiar with training, whereas their object selectivity increased (narrowed tuning) and generalized across degradation levels.

27. Sigala N, Logothetis NK: **Visual categorization shapes feature selectivity in the primate temporal cortex**. *Nature* 2002, **415**:318-320.

28. Rainer G, Miller EK: **Effects of visual experience on the representation of objects in the prefrontal cortex**. *Neuron* 2000, **27**:179-189.

29. Baker CI, Behrmann M, Olson CR: **Impact of learning on representation of parts and wholes in monkey inferotemporal cortex**. *Nat Neurosci* 2002, **5**:1210-1216.

30. Freedman DJ, Riesenhuber M, Poggio T, Miller EK: **A comparison of primate prefrontal and inferior temporal cortices during visual categorization**. *J Neurosci* 2003, **23**:5235-5246.

31. Yang T, Maunsell JH: **The effect of perceptual learning on**
•• **neuronal responses in monkey visual area V4**. *J Neurosci* 2004,
24:1617-1626.
This paper is the latest in a series of impressively thorough monkey
neurophysiological studies of perceptual learning effects along the ventral
stream (V1, V2 and V4) [21]. In these studies the authors systematically
examine a large range of potential single neuron effects of orientation
discrimination training. The overall conclusion is that, whereas all effects
in V1 and V2 are small (especially when compared with results in other
primary sensory areas), V4 neurons near the trained visual field location
and stimulus orientation show moderately sharper orientation selectivity
and better population discrimination performance.

32. Logothetis NK, Pauls J, Poggio T: **Shape representation in the**
inferior temporal cortex of monkeys. *Curr Biol* 1995, **5**:552-563.

33. Jagadeesh B, Chelazzi L, Mishkin M, Desimone R: **Learning**
increases stimulus salience in anterior inferior temporal
cortex of the macaque. *J Neurophysiol* 2001, **86**:290-303.

34. Lee TS, Yang CF, Romero RD, Mumford D: **Neural activity in early**
visual cortex reflects behavioral experience and higher-order
perceptual saliency. *Nat Neurosci* 2002, **5**:589-597.

35. van Turennout M, Ellmore T, Martin A: **Long-lasting cortical**
plasticity in the object naming system. *Nat Neurosci* 2000,
3:1329-1334.

36. Grill-Spector K, Kushnir T, Hendler T, Malach R: **The dynamics of**
object-selective activation correlate with recognition
performance in humans. *Nat Neurosci* 2000, **3**:837-843.

37. Schiltz C, Bodart JM, Dubois S, Dejardin S, Michel C, Roucoux A,
Crommelinck M, Orban GA: **Neuronal mechanisms of**
perceptual learning: changes in human brain activity with
training in orientation discrimination. *Neuroimage* 1999,
9:46-62.

38. Dolan RJ, Fink GR, Rolls E, Booth M, Holmes A, Frackowiak RS,
Friston KJ: **How the brain learns to see objects and faces in an**
impoverished context. *Nature* 1997, **389**:596-599.

39. Gauthier I, Tarr MJ, Anderson AW, Skudlarski P, Gore JC:
Activation of the middle fusiform 'face area' increases with
expertise in recognizing novel objects. *Nat Neurosci* 1999,
2:568-573.

40. Furmanski CS, Schluppeck D, Engel SA: **Learning strengthens**
• **the response of primary visual cortex to simple patterns**.
Curr Biol 2004, **14**:573-578.
This study provides an elegant example of the link between experience-
dependent behavioral improvement and neural plasticity in the primary
visual cortex. In particular, after training with oblique orientation patterns,
behavioral performance in the discrimination of these patterns improved
and fMRI responses in V1 increased at magnitudes similar to those for
cardinal orientation patterns before training.

41. Kourtzi Z, Betts LR, Sarkheil P, Welchman AE: **Distributed neural**
• **plasticity for shape learning in the human visual cortex**. *PLoS*
Biol 2005, **3**:e204.
The authors show that the human brain learns novel objects in complex
scenes by reorganizing shape processing across visual areas, while
taking advantage of natural image correlations and optimising neural
coding for the task context.

42. Edelman S: *Representation and Recognition in Vision*.
MIT Press; 1999

43. Riesenhuber M, Poggio T: **Hierarchical models of object**
recognition in cortex. *Nat Neurosci* 1999, **2**:1019-1025.

44. Ullman S: *High Level Vision*. MIT Press; 1996.

45. Barlow H: **The neuron doctrine in perception.**. In *The Cognitive*
Neurosciences. Edited by Gazzaniga. MIT Press; 1995:415-435.

46. Fukushima K: **Neocognitron: a self organizing neural network**
model for a mechanism of pattern recognition unaffected by
shift in position. *Biol Cybern* 1980, **36**:193-202.

47. Fiser J, Aslin RN: **Unsupervised statistical learning of higher-**
order spatial structures from visual scenes. *Psychol Sci* 2001,
12:499-504.

48. Chun MM: **Contextual cueing of visual attention**. *Trends Cogn*
Sci 2000, **4**:170-178.

49. Simoncelli EP: **Vision and the statistics of the visual**
environment. *Curr Opin Neurobiol* 2003, **13**:144-149.

50. Messinger A, Squire LR, Zola SM, Albright TD: **Neural correlates**
• **of knowledge: stable representation of stimulus associations**
across variations in behavioral performance. *Neuron* 2005,
48:359-371.
This study provides evidence for a novel class of inferotemporal neurons
that represent arbitrary associations between pairs of stimuli independent
of whether the animal chooses the correct remembered visual stimulus or
not. These findings suggest that knowledge in the context of memorized
associations is represented in the inferior temporal cortex regardless of its
role and weight in behavioral decisions.

51. Messinger A, Squire LR, Zola SM, Albright TD: **Neuronal**
representations of stimulus associations develop in the
temporal lobe during learning. *Proc Natl Acad Sci USA* 2001,
98:12239-12244.

52. Rosenthal O, Fusi S, Hochstein S: **Forming classes by stimulus**
frequency: behavior and theory. *Proc Natl Acad Sci USA* 2001,
98:4265-4270.

53. Freedman DJ, Riesenhuber M, Poggio T, Miller EK: **Experience-**
• **dependent sharpening of visual shape selectivity in inferior**
temporal cortex. *Cereb Cortex* 2005. In press.
This neurophysiological study provides evidence for enhanced perfor-
mance in object learning in the context of a categorization task linked to
improved neural selectivity in IT. Interestingly, these learning effects were
specific to the trained object orientation and were observed not only for
objects with which the monkeys were explicitly trained but also for stimuli
to which the monkeys were passively exposed. These findings suggest
that similar neural plasticity mechanisms that result in the sharpening of
neural responses might underlie learning through supervised training or
passive experience.

54. Wallis G, Bulthoff HH: **Effects of temporal association on**
recognition memory. *Proc Natl Acad Sci USA* 2001,
98:4800-4804.

55. Stringer SM, Perry G, Rolls ET, Proske JH: **Learning invariant**
object recognition in the visual system with continuous
transformations. *Biol Cybern* 2006, **94**:128-142.

56. DiCarlo JJ, Maunsell JHR: **Anterior inferotemporal neurons of**
monkeys engaged in object recognition can be highly
sensitive to object retinal position. *J Neurophysiol* 2003,
89:3264-3278.

57. Kourtzi Z, Shiffrar M: **One-shot view invariance in a moving**
world. *Psychol Sci* 1997, **8**:461-466.

58. Cox DD, Meier P, Oertelt N, DiCarlo JJ: **'Breaking' position-**
• **invariant object recognition**. *Nat Neurosci* 2005, **8**:1145-1147.
In normal, real-world viewing, the same object is seen at different posi-
tions across short time intervals as the eyes rapidly explore the visual
world. To test the idea that this experience might underlie the position-
tolerance of recognition, the authors introduced undetected changes in
object identity during eye movements. This manipulation disrupted the
position tolerance of object recognition in healthy adult observers. This
shows that position tolerance is plastic in adults, and suggests that even
this bedrock property of primate object recognition might be acquired or
maintained by natural visual experience.

59. Wang G, Obama S, Yamashita W, Sugihara T, Tanaka K:
Prior experience of rotation is not required for recognizing
objects seen from different angles. *Nat Neurosci* 2005,
8:1768-1775.

60. Rolls ET, Aggelopoulos NC, Zheng F: **The receptive fields of**
inferior temporal cortex neurons in natural scenes. *J Neurosci*
2003, **23**:339-348.

61. Sheinberg DL, Logothetis NK: **Noticing familiar objects in real**
world scenes: the role of temporal cortical neurons in natural
vision. *J Neurosci* 2001, **21**:1340-1350.

62. Gold J, Bennett PJ, Sekuler AB: **Signal but not noise changes**
with perceptual learning. *Nature* 1999, **402**:176-178.

63. Brady MJ, Kersten D: **Bootstrapped learning of novel objects**.
J Vis 2003, **3**:413-422.

64. Dosher BA, Lu ZL: **Perceptual learning reflects external noise**
filtering and internal noise reduction through channel
reweighting. *Proc Natl Acad Sci USA* 1998, **95**:13988-13993.

65. Li RW, Levi DM, Klein SA: **Perceptual learning improves efficiency by re-tuning the decision 'template' for position discrimination**. *Nat Neurosci* 2004, **7**:178-183.

66. Eckstein MP, Abbey CK, Pham BT, Shimozaki SS: **Perceptual**
• **learning through optimization of attentional weighting: human versus optimal bayesian learner**. *J Vis* 2004, **4**:1006-1019.
This paper introduces a new experimental paradigm into the study of perceptual learning by comparing human and optimal Bayesian learners. The results suggest that humans learn to localize targets with uncertainty about orientation and polarity within a brief sequence of trials, but they rely more on previous decisions than on feedback, resulting in slower and more incomplete learning than that of an ideal observer.

67. Barlow H: **Conditions for versatile learning, Helmholtz's unconscious inference, and the task of perception**. *Vision Res* 1990, **30**:1561-1571.

68. Cox D, Meyers E, Sinha P: **Contextually evoked object-specific responses in human visual cortex**. *Science* 2004, **304**:115-117.

69. Bar M, Aminoff E: **Cortical analysis of visual context**. *Neuron* 2003, **38**:347-358.

70. Zoccolan D, Cox DD, DiCarlo JJ: **Multiple object response normalization in monkey inferotemporal cortex**. *J Neurosci* 2005, **25**:8150-8164.

71. Aggelopoulos NC, Rolls ET: **Scene perception: inferior temporal cortex neurons encode the positions of different objects in the scene**. *Eur J Neurosci* 2005, **22**:2903-2916.

72. Crist RE, Li W, Gilbert CD: **Learning to see: experience and attention in primary visual cortex**. *Nat Neurosci* 2001, **4**:519-525.

73. Schoups AA, Vogels R, Orban GA: **Human perceptual learning in identifying the oblique orientation: retinotopy, orientation specificity and monocularity**. *J Physiol* 1995, **483**:797-810.

74. Schwartz S, Maquet P, Frith C: **Neural correlates of perceptual learning: a functional MRI study of visual texture discrimination**. *Proc Natl Acad Sci USA* 2002, **99**:17137-17142.

75. Crist RE, Li W, Gilbert CD: **Learning to see: experience and attention in primary visual cortex**. *Nat Neurosci* 2001, **4**:519-525.

76. Sigman M, Pan H, Yang Y, Stern E, Silbersweig D, Gilbert CD: **Top-**
• **down reorganization of activity in the visual pathway after learning a shape identification task**. *Neuron* 2005, **46**:823-835.
The authors report large-scale reorganization across cortical networks for task-dependent learning. Learning to detect a target among distractors in the context of a visual search task resulted in increased activations in retinotopic areas By contrast, decreased activations were observed both in higher occipitotemporal areas involved in shape analysis and in frontal and parietal areas involved in attentional processing.

77. Ahissar M, Hochstein S: **The reverse hierarchy theory of visual**
•• **perceptual learning**. *Trends Cogn Sci* 2004, **8**:457-464.
This review highlights psychophysical and physiological evidence that learning begins at higher stages of visual analysis and proceeds in a top-down manner to earlier stages when finer analysis of the stimulus is necessary.

78. Buchel C, Coull JT, Friston KJ: **The predictive value of changes in effective connectivity for human learning**. *Science* 1999, **283**:1538-1541.

79. McIntosh AR, Rajah MN, Lobaugh NJ: **Interactions of prefrontal cortex in relation to awareness in sensory learning**. *Science* 1999, **284**:1531-1533.

80. Li W, Piech V, Gilbert CD: **Perceptual learning and top-down influences in primary visual cortex**. *Nat Neurosci* 2004, **7**:651-657.

81. Sagi D, Tanne D: **Perceptual learning: learning to see**. *Curr Opin Neurobiol* 1994, **4**:195-199.

82. Sigman M, Gilbert CD: **Learning to find a shape**. *Nat Neurosci* 2000, **3**:264-269.