PERCEPTUAL LEARNING OF POP-OUT
AND THE PRIMARY VISUAL CORTEX

LI ZHAOPING

University College London, United Kingdom

I propose that perceptual learning of tasks to detect targets among uniform background items involves changing intra-cortical interactions in the primary visual cortex (V1). This is the case for tasks that rely mainly on bottom-up saliency to guide attention to the task relevant locations quickly, and rely less on top-down knowledge of the stimuli or on other strategies. In particular, suppression between V1 neurons responding to background, rather than target, visual items is predicted to increase over the course of such learning. Various other predictions are derived from this proposal, based on the theory that V1 creates a bottom-up saliency map to guide attention. Different tasks depend to different degrees on attention driven by bottom-up saliency; this leads to differences among findings from various studies of perceptual learning of pop out or detection tasks.

Keywords: saliency; V1; intra-cortical connections; attention; specificity of learning; visual search

INTRODUCTION

Studies into perceptual learning of pop out in visual search or segmentation have led to diverse results. For instance, Karni and Sagi (1991) found great specificity in learning to segment a texture of oriented bars from a background texture in a brief display (see *Fig. 1A*), presumably utilizing pop-out signals from the foreground texture. The effect of learning when stimuli were seen in one eye does not transfer to the other eye; equally, when the foreground texture was in one part of the visual field during learning, transfer did not occur to other parts. Learning was also specific for the orientation of the texture bars in the background texture.
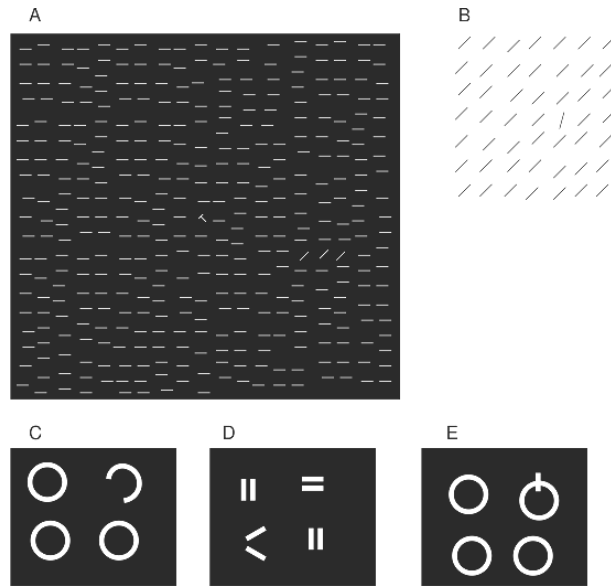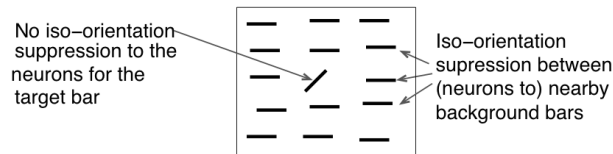
Figure 1. Stimuli patterned after those in the experiments investigating perceptual learning of pop-out in visual search or segmentation. A: a stimulus as used by Karni and Sagi (1991) and Schoups and Orban (1996). Whilst executing a demanding fixation task, the task of interest is to indicate, based on a brief display, whether the foreground texture, whose position could vary (within a quadrant of the background texture) between trials, had three horizontally or vertically arranged oblique bars. The foreground texture is at least 3 bars away from the center of a 19 × 19 texture array. B: a stimulus as used by Ahissar and Hochstein (1996). Their task involved detecting an orientation singleton. The stimulus was briefly displayed, and the target was typically only 1–2 texture bars away from the fixation point. C, D, E: visual search stimuli like those used by Sireteanu and Rettenbach (1996). Subjects had to report whether there was an odd-one-out item. There were only 1, 4, 8, 16 items in the display, each 3.5° in size, and the stimulus was displayed until the subjects responded.

Since the primary visual cortex has neurons tuned to orientations of bars, and is the only cortical area with a substantial number of neurons tuned to ocular origin, Karni and Sagi (1993) suggested that the learning involved changes in the primary visual cortex. However, puzzlingly, they also found that the learning is not specific to the orientation of the foreground texture bars. Their experiments were repeated with similar stimuli (but a different mask, see later) and procedures by Schoups and Orban (1996) who, by contrast, found complete ocular transfer of the learning. Meanwhile, experiments by Ahissar and Hochstein (1996) found that learning to detect an orientation singleton in an uniform texture of bars (see *Fig. 1B*) was not specific to ocular origin for most subjects, but was specific for the orientation of both the background and target orientation. Hence, they suggested that the brain site responsible for the learning is somewhere that receives inputs from the primary visual cortex. Furthermore, Sireteanu and Rettenbach (1996) found that practice can quickly make easier some initially difficult visual searches for an "odd one out" target (see *Fig. 1C, D, E*), and that the effect of learning is unspecific not only to the eye of origin, but also to the target and non-target item
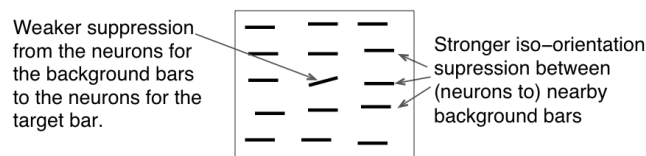
types. They suggested this implicates a much more central site of learning than the primary visual cortex.

Assuming that all the studies are indeed investigating the learning of pop-out, the diversity of findings about specificity is puzzling, and makes pin-pointing the neural basis of such learning hard. One issue is time-scale. It has been suggested (Karni and Sagi, 1993) that learning on the fast time-scale of a few minutes is more likely to involve higher brain areas than learning on a slower time scale of days. Hence, it is possible that different durations of learning in different studies contribute to this diversity. For instance, learning in the study by Sireteanu and Rettenbach took a much shorter time than those in the other studies. This paper does not focus on this factor of learning time-scales. Instead I propose a neural mechanism in the primary visual cortex for the perceptual learning of pop out. This is motivated by recent developments indicating that the primary visual cortex (V1) is the neural basis for computing bottom-up visual saliency and thus guiding attention to the pop out location. This suggests that perceptual learning to detect or segment some visual inputs is likely to involve changes in V1, provided the task performance relies mainly on bottom-up or stimulus driven saliency to detect or locate the relevant visual input. As I will suggest in detail in this paper, one of the main reasons for the divergent findings is that the tasks in different studies differ in the extent to which they rely on stimulus-driven saliency signals.

A: the target bar and the background bars have very different orientations

No iso–orientation suppression to the neurons for the target bar

Iso–orientation supression between (neurons to) nearby background bars

B: the target bar and the background bars have very similar orientations

Weaker suppression from the neurons for the background bars to the neurons for the target bar.

Stronger iso–orientation supression between (neurons to) nearby background bars

C: a few target bars

Iso–orientation suppression is weaker, since there are fewer iso–oriented neighbors.

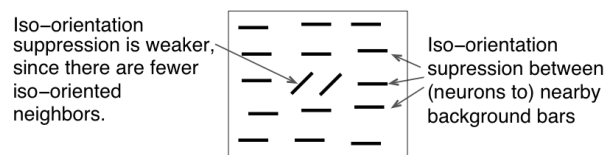Iso–orientation supression between (neurons to) nearby background bars

Figure 2. Three (A, B, C) different stimulus conditions for target detection, and the relevant neural interactions involved. In each case, V1 neurons responding to the background bars are suppressed by iso-orientation suppression to each other. Responses to the target bar(s) is highest among neurons responding to the scene, since suppression to the neurons responding to the target bar(s) is absent in A, and weaker in B and C, making the target bar(s) more salient. Learning the target detection is proposed to involve changes in the strengths of the various suppressions.

Based on physiological observations of contextual influences in V1 such as iso-orientation suppression (Allman, Miezin and McGuinness, 1985; Knierim and van Essen, 1992; Li and Li, 1994; Lamme, 1995; Sillito et al., 1995; Kastner, Nothdurft and Pigarev, 1997, 1999; Nothdurft, Gallant and van Essen, 1999, Jones et al. 2001; Wachtler, Sejnowski and Albright, 2003), and anatomical observations of intra-cortical connections linking nearby neurons tuned to similar features (Rockland and Lund, 1983; Gilbert and Wiesel, 1983; Bosking et al., 1997), Li (1999a, 2002) proposed that the relative V1 responses to different locations in a scene correspond to the relative saliencies of those locations, with the most salient location in a scene being the receptive field location of the V1 neuron that is most activated to the scene.

The mechanism in V1 that transforms its input visual contrast signals to its output saliency signals involves intra-cortical interactions between neurons that mediate contextual influences such as iso-feature suppression. Thus, as it should be, saliency is context dependent. This was demonstrated in a V1 model (Li, 1999b, 2000).

In particular, as illustrated in *Figure 2,* when there is a background texture of uniformly oriented bars, as in the stimuli used in the perceptual learning experiments (Karni and Sagi 1991; Schoups and Orban 1996; Ahissar and Hochstein 1996), nearby V1 neurons responding to these iso-oriented bars will suppress each other via di-synaptic inhibitory connections. The neurons responding to the foreground or target texture bars are not suppressed by neurons responding to the background bars when the foreground texture bars have a sufficiently different orientation from that of the background bars. This is because the di-synaptic connections tend to link neurons tuned to similar (orientation) features. Thus, in the experiment of Ahissar and Hochstein (1996), the orientation singleton foreground bar, like in *Fig. 2A,* should evoke the highest response in the scene, and it should thus be the most salient according to the V1 theory.

Even when there is more than one foreground bar, as in *Fig. 2C,* provided that there are not too many, the foreground bars will still be the most salient. This is because the iso-orientation suppression between neurons responding to the different foreground bars is much weaker than the iso-orientation suppression between the neurons responding to the much more numerous background bars. This is true of the stimuli used by Karni and Sagi (1991, *Fig. 1A*).

The behavioral predictions from this V1 theory of saliency have been experimentally tested and confirmed (Zhaoping and May, 2007; Koene and Zhaoping, 2007), including a surprising prediction that an eye-of-origin singleton can capture attention automatically even though it is not distinctive to awareness (Zhaoping, 2008).

In this paper, I propose that in perceptual learning paradigms that involve increases in the strength of the bottom-up pop-out, the intra-cortical or recurrent connections in V1 linking neurons tuned to the same or similar features will change. In particular, to enhance bottom-up pop out of an orientation singleton or a border between textures of oriented bars, perceptual learning is proposed to involve changes in the strength of di-synaptic inhibition between nearby V1 neurons tuned to the same or similar orientations.

For instance, for the case of the experiments by Karni and Sagi (1991), even though the foreground texture bars should be the most salient, they may not be sufficiently salient to attract attention adequately for the task when viewed for the very short duration employed in their experiments. Task performance could be facilitated when the relative saliency of the foreground against the background bars is increased, through perceptual learning.

According to the V1 saliency mechanism described above, there are actually two ways to enhance relative foreground saliency: one, as mentioned, is to increase the iso-orientation suppression between the V1 neurons responding to the background bars, making these bars even less salient; the other is to increase the responses of the V1 neurons to the foreground bars by (1) increasing the sensitivities of these neurons, (2) reducing the iso-orientation suppression between these neurons, and/or (3) decreasing the suppression from the neurons responding to the background bars to neurons responding to the target bar(s) when they are oriented sufficiently similar to each other, as in *Figure 2B,* since intra-cortical connections also link neurons tuned to similar but not identical features although such connections are presumably scarcer and/or weaker than the iso-feature connections. The possibility (1) is not favoured by the dearth of evidence that perceptual learning changes the sensitivities of V1 neurons to isolated stimuli (without context). This leaves changes in intra-cortical suppression, between neurons responding to the background bar, or between neurons responding to the foreground bars, or from neurons responding to the background bars to other neurons to the foreground bars when the bars have sufficiently similar orientations or features, or a combination of these, as the mechanism. I omit the possibility (3) here assuming that the feature (orientation) of the target bar(s) is sufficiently different from that of the background bars unless stated otherwise. As there are many more background than foreground bars, it may be more effective to increase the iso-orientation suppression between the background bars to facilitate pop-out. However, whether this proposed mechanism for enhancing pop out is actually involved in the perceptual learning of any visual search or discrimination task will depend on whether the learned task relies mainly on the bottom-up saliency signal, or on other signals and strategies. The discrepancies between the specificities of learning in the previous studies are likely to arise from the differences in the effective signals and strategies for the task, as detailed below.

PREDICTIONS AND RELATION TO THE PREVIOUS OBSERVATIONS

If, as proposed, increasing iso-orientation suppression between the V1 neurons responding to the background texture bars is the dominant mechanism for perceptual learning to detect orientation singleton or target texture, then various predictions follow.

Firstly, the learning should be specific to the orientation of the background texture bars and not to the orientation of the foreground texture bars. In particular, using *Figure 2A* for illustration, when the suppression between the background horizontal bars is increased, the relative saliency of the foreground bar is independent of its orientation as long as it is oriented sufficiently differently from the background bars. So if training proceeds for an oblique target bar tilted clock-wise from horizontal background bars, the effect of learning should transfer to the same task with an oblique target bar tilted counter-clock-wise from the same set of the horizontal background bars. However, if the orientation of the background bars is changed, then since the iso-orientation suppression between the neurons to the new set of background bars has not been changed through learning, the task performance should drop back to the pre-learning level. This specificity to the background but not foreground bars, though seemingly puzzling at first sight, is indeed observed by Karni and Sagi (1991).

The specificity is, however, not observed by Ahissar and Hochstein in their case of learning to detect orientation singletons. A likely reason is that the target bar in their stimulus is too close to the fixation point, mostly about only 1 or 2 texture elements away. According to observations by Motter and Belky (1998), this target is mainly within the attentional window centered at the display center (the fixation point). The saliency signal, whose putative purpose is mainly to guide attention to the task relevant location outside the current attentional window, is then redundant or irrelevant for the task. Their perceptual learning is likely to involve improving other strategies or signals, perhaps to facilitate the discrimination between the target-less stimulus from ones with a near-fovea target when all these stimuli are within the attentional window. As the V1 saliency theory only applies to bottom-up attentional guidance, it is thus less likely to apply to the findings by Ahissar and Hochstein.

In contrast, the foreground texture in Karni and Sagi's experiment is at a greater distance from the attentional window centered at fixation. Furthermore, the subjects in their experiments were forced to focus their attention at fixation because they also had to address a demanding letter discrimination task. This would limit their ability to diffuse an attentional focus within the brief stimulus presentation time. Thus, given the uncertainty in the location of the foreground bars, the limited attentional resource, and the brief display, the subjects must rely mainly on saliency signals to locate the foreground bars (which are outside their attentional window), before they can discriminate whether these bars are arranged vertically or horizontally.

Secondly, the proposal predicts that, if the displayed items are insufficiently dense or numerous, it is less likely for the task to involve bottom-up saliency computed by V1. Therefore, the learning would not be as likely to be specific in terms of low level features such as orientation and eye of origin. This is because the intra-cortical connections between V1 neurons are of finite range, only about 2–4 millimeters, and thus only neurons responding to sufficiently nearby input items can interact with each other to achieve the iso-feature suppression necessary for singleton pop-out. When the items are too sparse or too far away from each other, or are too big, the iso-feature suppression between the background items would be too weak or non-existent, making the singleton or odd-one-out item less salient. It is indeed known that a feature singleton pops out more readily from a denser background texture (of uniform items, as long as it is not too dense, Nothdurft, 2000). When the bottom-up V1-based saliency signal is insufficient, because a display is too sparse, visual search for the odd-one-out has to be carried out using other strategies. If these strategies involve higher level processes, it would imply that the learning would exhibit less specificity associated with low level features. Indeed, stimuli in Sireteanu and Rettenbach's experiments contain only 1, 4, 8, and 16 items, much sparser than those used by Karni and Sagi (1991). The observed effects of learning are indeed not so specific even to the stimulus types.

Thirdly, the proposal predicts that there should be no specificity to the target location among the background items. This is because the target should be relatively salient against all background bars in order to attract bottom-up attention effectively – this is of course, assuming that there is no top-down knowledge regarding the likely locations of the target (allowing impossible target locations to be ignored). Hence, iso-orientation suppression between the background bars should be increased between all nearby background bars, regardless of their distance from the target bar, as long as the subject does not exclude from prior knowledge any

background location as a possible target location. Hence, if learning used stimuli with targets at one subset of locations in the background texture, the effect of learning should transfer to targets at the other locations.

This prediction apparently disagrees with findings in all the three perceptual learning studies involving orientation detection or segmentation (Karni and Sagi, 1991; Schoups and Orban, 1996; Ahissar and Hochstein, 1996). The most likely explanation for this discrepancy is that, in these experiments, when the target is constrained within a limited spatial region, e.g. a particular quadrant, subjects can use top-down attentional control to avoid their attention being distracted to the other quadrants where the target is known not to appear. This means, it is unnecessary for the learning to increase the iso-orientation suppression between neurons responding to the background bars in the task irrelevant quadrants. However, this artificially introduces specificity to target location – when the target appears in the new quadrants among background bars whose mutual suppression has not been strengthened, the target saliency should revert back to the pre-learning level. In order to make subjects unaware of the restricted target locations during training so that they cannot use top-down guidance, it is best to interleave the training target positions spatially with respect to the untrained, testing, positions, e.g. having the two sets make a checkerboard pattern in space, with positions of different sets occupying different types of checks. This way, when the training positions are sufficiently numerous, subjects may believe that the target could appear anywhere within the background texture. Note that, learning should still be specific to the location of the whole stimulus pattern, e.g. if learning used the whole stimulus presented to the left half of the visual field, it should not transfer to stimulus in the right half of the visual field.

Fourthly, the proposal predicts that whether or not learning will be specific to the eye of origin depends on the precise stimulus and the adaptation conditions used in the experiments. This is because some V1 neurons are monocular and tuned to one particular eye of origin; whereas others are binocular and not tuned to eyes of origin. If intra-cortical inhibition between binocular neurons is changed during learning when the stimuli are viewed by one eye only, then the effects of learning should transfer to the other eye; if the changed inhibition is between the monocular neurons, then there should be little transfer. There is evidence that V1 neurons with larger receptive fields are more tuned to ocular origin than those with smaller receptive fields (Poggio and Fischer, 1977, and Anzai, private communication 2006), and this is also predicted from a theory of efficient stereo coding (Li and Atick, 1994; Li, 1995). The current proposal predicts that pop-out learning using textures of larger elements (but not too much larger than the receptive fields of the monocular neurons) is more likely to involve monocular neurons, making eye of origin specificity more likely.

Karni and Sagi (1991) observed ocular specificity, while the other authors mentioned above did not. The texture elements used in different experiments are not of the same size, and this could be one reason. In particular, Ahissar and Hochstein's texture had smaller elements. Schoups and Orban (1996) used almost the same stimuli as Karni and Sagi (1991), but their texture elements were slightly thinner and longer, and observed no such specificity. More importantly, according to the theory of efficient stereo coding (Li and Atick, 1994; Li, 1995), V1 neurons' ocular selectivity should quickly (in minutes) adapt to the binocular correlation in the visual environment. In the natural visual environment, the inputs to the two eyes are correlated to some degree. However, more V1 neurons will become monocular when the visual in-

puts to the two eyes become less correlated through experimental manipulations. There has been some preliminary experimental evidence supporting this predicted change in neuron's ocular tuning (Chen and Li, 1998). Specifically, if observers viewed the world monocularly, one eye at a time, their binocular inputs will be less correlated than usual, and this should induce adaptation leading to more monocular cells in V1. Hence, if Karni and Sagi (1991) let the subjects view the environment monocularly, e.g. by wearing monocular goggles, for a longer time before and during their experiments than Schoups and Orban (1996), it would account for the puzzling disagreement between their observations as to this specificity.

To summarize, strong ocular transfer may not invalidate V1 as the substrate for learning. Furthermore, ocular specificity in learning is more easily induced using larger texture elements and letting subjects adapt to a monocular or ocularly less correlated visual environment. Meanwhile, the observation of ocular specificity, even though observed only by Karni and Sagi (1991), supports the idea that V1 is a plausible substrate for learning pop-out.

In some situations, the learning of bottom-up pop out is predicted to involve changes in V1 other than increasing iso-orientation suppression between neurons responding to the background bars. If the target and background bars have sufficiently similar orientations, as in *Figure 2B,* V1 neurons responding to the background bars can also suppress those responding to the target bar, since the intra-cortical connections tend also to link neurons tuned to similar and not necessarily identical orientations. In such a case, pop out can be enhanced by learning to reduce the intra-cortical suppression from the neurons responding to the background bars to the neurons responding to the target bar, but not vice versa, so that the responses to the target but not to the background bars are less suppressed. When there are more than one target bar to form a foreground texture as in *Figure 2C,* learning can also involve reducing iso-orientation suppression between V1 neurons responding to different target bars, again to reduce suppression of responses to the targets. Both situations above should evoke a specificity of learning to the target orientation, such that the transfer to a different target orientation will not be 100%. The second situation above should also weaken the strong specificity of learning to the orientation of the background bars, such that transfer to a different orientation of the background bars will not be completely zero.

DISCUSSION

In summary, it is proposed that under certain stimulus, task, and adaptation conditions, perceptual learning of visual search and texture segmentation should involve a change in the intra-cortical interactions in V1 associated with iso-feature suppression, thus strengthening bottom-up saliency of the task relevant locations. These conditions are: (1) the search target is sufficiently far in the periphery and outside the attentional window centered at fixation, so that bottom-up saliency plays a key role in guiding attention to the task relevant location; (2) successful performance of the task requires quick attentional guidance to the task relevant location; (3) the distractor items are sufficiently dense and identical to each other. This implies that neurons activated by different distractors can be linked by finite range intra-cortical V1 connections which mediate iso-feature suppression. It is these connections that are changed over the course of learning to make the distractors less salient; (4) subjects have limited prior

knowledge of the possible target locations, so that they cannot use top-down attentional guidance to the target and have to rely heavily on bottom-up saliency.

Once these conditions are satisfied, signatures of V1's involvement should be more apparent. These signatures are predicted to be: specificity to the background features but not the target/foreground features when the foreground has no more than one or a few elements whose features are sufficiently different from those of the background elements, specificity to the eye of origin when stimulus and adaptation condition encourages involvement of monocular neurons (see above), and a lack of specificity to the target location if the subjects are restricted from having or utilizing the knowledge of possible target locations. It should be noted that, contrary to conventional expectations, an involvement of V1 mechanisms predicts a lack of specificities to features that are coded by V1 neurons, such as target orientation and location. Note that while analysis in this paper mainly refers to the orientation feature for target detection, the arguments and conclusions generalize readily for target detection by other V1 features such as color and motion direction.

We also suggest that the degrees to which the predicted specificities of learning differ from the observations in the previous studies are determined by the extent to which the stimuli, tasks, and adaptation conditions differ from the required conditions listed above. For the same reason, the divergent observations from different studies are also the results of their different stimuli, tasks, and adaptation conditions. Since the effect of saliency typically increases with eccentricity and decreases with top-down knowledge, signatures of V1's involvement in the learning should be stronger and more apparent if the foreground texture is more eccentric, in a larger array of background texture, and has a larger degree of uncertainty regarding its possible location (even when the target location is within a single quadrant of the display). Also, since the role of bottom-up saliency is stronger in tasks demanding quicker attentional guidance, removing the task relevant stimulus quickly and completely from perception using an effective mask should enhance the role of bottom-up saliency in the task. In this regard, it is noteworthy that the mask used by Karni and Sagi (1991) differs from those used by Schoups and Orban (1996) and Ahissar and Hochstein (1996) – it is more irregular and is thus more effective in removing the target stimulus from perception, as it is known that target stimulus can more easily shine through regular masks or contexts presented immediately after it (Herzog and Fahle 2002).

Changes of intra-cortical interactions in V1 through perceptual learning have been experimentally observed (Crist, Li and Gilbert, 2001; Li, Piëch and Gilbert, 2004) for discrimination of stimulus at a known location. By analyzing the task to detect a target from a background of uniformly oriented bars, this paper arrives at specific predictions of the changes in the intra-cortical suppression from a theory of V1's role in bottom-up saliency. Experimental tests of these predictions will be informative and desirable.

As human decision-making can be affected by many factors, a typical visual search task will recruit many types of signals and involve many possible decision-making strategies. Some signals are bottom-up or stimulus driven in nature; others are more top-down and influenced by bias, knowledge, and other factors. Different decision-making strategies can be optimal to different degrees, and it is not uncommon for observers to use sub-optimal strategies first, before improving them. For instance, we have shown that subjects can start out using task-irrelevant information (in our case, the rotational-invariance of object detection) even

though, by design, this was detrimental to visual search performance (Zhaoping and Guyader, 2007). This sub-optimal strategy can be corrected very quickly through learning. Indeed, this strategy learning process is typically faster (and has a profound effect on the reaction time of the task performance) than learning to enhance the task relevant saliency signal from the orientation singleton in the target.

Given the involvement of so many factors in decision-making for any particular task, learning could occur at multiple levels, and involve different processes operating at different time scales (Karni and Sagi, 1993). It is thus not unexpected to see seemingly different observations for very similar tasks. Indeed, it is not easy to disentangle the multiple mechanisms involved in generating the net effects of multiple processes. This is especially the case when learning involves higher level or top-down processes which can be very subject dependent, and thus may not be completely reproducible across subjects. All these complications make it difficult to study mechanisms of learning. However, once we have a candidate mechanism for learning, such as the V1 mechanism proposed here, and once we understand the limits of its applicability and conditions for its manifestation, we can design experiments accordingly to better isolate or highlight the candidate mechanism in order to dissect, diagnose, and test hypotheses about this one single mechanism more effectively. This approach can hopefully help us to probe different mechanisms one at a time.

ACKNOWLEDGEMENTS

REFERENCES

Allman, J., Miezin, F., McGuinness, E. L. (1985): Stimulus specific responses from beyond the classical receptive field: Neurophysiological mechanisms for local-global comparison in visual neurons. *Annu. Rev. Neurosci., 8,* 407–430.

Ahissar, M., Hochstein, S. (1996): Learning pop-out detection: Specificities to stimulus characteristics. *Vision Res., 36,* 3487–3500.

Bosking, W. H., Zhang, Y., Schofield, B., Fitzpatrick, D. (1997): Orientation selectivity and the arrangementof horizontal connections in tree shrew striate cortex. *J. Neurosci., 17*(6), 2112–2127.

Chen, D., Li, Z. (1998): A psychophysical experiment to test the efficient stereo coding theory. In: Wong, K. M., King, I., Yeung, D. Y. (eds): *Theoretical Aspects of Neural Computation*. Springer-Verlag, pp. 225–235.

Crist, R. T., Li, W., Gilbert, C. D. (2001): Learning to see: experience and attention in primary visual cortex. *Nature Neuroscience, 4,* 519–525.

Gilbert, C. D., Wiesel, T. N. (1983): Clustered intrinsic connections in cat visual cortex. *J. Neurosci., 3*(5), 1116–1133.

Herzog, M. H., Fahle, M. (2002): Effects of grouping in contextual modulation. *Nature, 415,* 433–436.

Jones, H. E., Grieve, K. L., Wang, W., Sillito, A. M. (2001): Surround suppression in primate V1. *J. Neurophysiol., 86*(4), 2011–2028.

Karni, A., Sagi, D. (1991): Where practice makes perfect in texture discrimination: Evidence for primary visual cortex plasticity. *Proc. Natl. Acad. Sci. USA, 88,* 4966–4970.

Karni, A., Sagi, D. (1993): The time course of learning a visual skill. *Nature, 365,* 250–252.

Kastner, S., Nothdurft, H. C., Pigarev, I. N. (1997): Neuronal correlates of pop-out in cat striate cortex. *Vision Research, 37,* 371–376.

Kastner, S., Nothdurft, H. C., Pigarev, I. N. (1999): Neuronal responses to orientation and motion contrast in cat striate cortex. *Visual Neuroscience, 16,* 587–600.

Knierim, J. J., van Essen, D. C. (1992): Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *Journal of Neurophysiology, 67,* 961–980.

Koene, A. R., Zhaoping, L. (2007): Feature-specific interactions in salience from combined feature contrasts: Evidence for a bottom-up saliency map in V1. *Journal of Vision, 7*, 1–14.

Lamme, V. A. F. (1995): The neurophysiology of figure-ground segregation in primary visual cortex. *Journal of Neuroscience, 15,* 1605–1615.

Li, C. Y., Li, W. (1994): Extensive integration field beyond the classical receptive field of cat's striate cortical neurons – classification and tuning properties. *Vision Research, 34*(18), 2337–2355.

Li, W., Piëch, V., Gilbert, C. D. (2004): Perceptual learning and top-down influences in primary visual cortex. *Nature Neuroscience, 7,* 651–657.

Li, Z. (1995): Understanding ocular dominance development from binocular input statistics. In: Bower, J. (ed.): *The Neurobiology of Computation.* Proceeding of Computational Neuroscience Conference 1994. Kluwer Academic Publishers, pp. 397–402.

Li, Z. (1999a): Contextual influences in V1 as a basis for pop out and asymmetry in visual search. *Proc. Natl. Acad. Sci. USA, 96*(18), 10530–10535.

Li, Z. (1999b): Visual segmentation by contextual influences via intra-cortical interactions in the primary visual cortex. *Network, 10*(2), 187–212

Li, Z. (2000): Pre-attentive segmentation in the primary visual cortex. *Spatial Vision, 13*(1), 25–50.

Li, Z. (2002): A saliency map in primary visual cortex. *TRENDS in Cognitive Sciences, 6*(1), 9–16.

Li, Z., Atick, J. J. (1994): Efficient stereo coding in the multiscale representation. *Network: Computations in Neural Systems, 5,* 157–174.

Motter, B. C., Belky, E. J. (1998): The zone of focal attention during active visual search. *Vision Research 38,* 1007–1022.

Nothdurft, H. C. (2000): Salience from feature contrast: Variations with texture density. *Vision Research, 40,* 3181–3200.

Nothdurft, H. C., Gallant, J. L., van Essen, D. C. (1999): Response modulation by texture surround in primate area V1: Correlates of 'popout' under anesthesia. *Visual Neuroscience, 16,* 15–34.

Poggio, G. F., Fischer, B. (1977): Binocular interaction and depth sensitivivty in striate and prestriate cortex of behaing rhesus monkey. *Journal of Neurophysiology, 40,* 1392–1405.

Rockland, K. S., Lund, J. S. (1983): Intrinsic laminar lattice connections in primate visual cortex. *J. Comp. Neurol., 216*(3), 303–318.

Schoups, A., Orban, G. (1996): Interocular transfer in perceptual learning of a pop-out discrimination task. *Proc. Natl. Acad. Sci. USA, 93,* 7358–7362.

Sillito, A. M., Grieve, K. L., Jones, H. E., Cudeiro, J., Davis, J. (1995): Visual cortical mechanisms detecting focal orientation discontinuities. *Nature, 378,* 492–496.

Sireteanu, R., Rettenbach, R. (1996): Perceptual learning in visual search: Fast, enduring, but non-specific. *Vision Research, 35,* 2037–2043.

Wachtler, T., Sejnowski, T. J., Albright, T. D. (2003): Representation fo color stimuli in awake macaque primary visual cortex. *Neuron, 37*(4), 681–691.

Zhaoping, L. (2008): Attention capture by eye of origin singletons even without awareness – a hallmark of bottom-up saliency map in the primary visual cortex. *Journal of Vision, 8*, 1–18.

Zhaoping, L., Guyader, N. (2007): Interference with bottom-up feature detection by higher-level object recognition. *Curr. Biol., 17*(1), 26–31.

Zhaoping, L., May, K. A. (2007): Psychophysical tests of the hypothesis of a bottom-up saliency map in primary visual cortex. *Public Library of Science, Computational Biology, 3*(4), e62.