The Role of Feedback in the Determination of Figure and Ground: A Combined Behavioral & Modeling Study

Lawrence A. Watling (l.watling@psychology.bbk.ac.uk)
Centre for Brain & Cognitive Development, School of Psychology, Birkbeck College, London, U.K.

Michael W. Spratling (michael.spratling@kcl.ac.uk)
Division of Engineering, King's College London, U.K.

Kris De Meyer (kris.de_meyer@ kcl.ac.uk)
Division of Engineering, King's College London, U.K.

Mark H. Johnson (mark.johnson@bbk.ac.uk)
Centre for Brain & Cognitive Development, School of Psychology, Birkbeck College, London, U.K.

Abstract

Object knowledge can exert on important influence on even the earliest stages of visual processing. This study demonstrates how a familiarity bias, acquired only briefly before testing, can affect the subsequent segmentation of an otherwise ambiguous figure-ground array, in favor of perceiving the familiar shape as figure. The behavioral data are then replicated using a biologically plausible neural network model that employs feedback connections to implement the demonstrated familiarity bias.

Keywords: figure-ground segmentation; top-down processing; cortical feedback; cortical visual system; neural networks.

Introduction

The visual system is made up of a complex network of interconnected neural processing units that communicate with one another across synapses that have varying efficacies, that are constantly subject to change, and that represent the associations between units. This network is made up of a series of cortical regions that are connected to create a processing hierarchy (Lamme & Roelfsema, 2000). The distribution of activity within the network represents the inputs, intentions and perceptions of the system. The primary visual cortex (V1) is the lowest level in this cortical hierarchy and is made up of cells with very small receptive fields that are selective to elementary features such as contour segments at specific orientations (Field et al., 1993; Gilbert et al., 1996; Kapadia et al., 1995). As the stimulus information ascends through the visual hierarchy it is progressively grouped, so that at each level the neurons receive input from multiple lower layer neurons and hence have progressively larger receptive fields. This grouping enables the neurons in each successive level to become selective to combinations of more and more features, until eventually the neurons at the top of this hierarchy (in the temporal cortex) are tuned to extremely complex stimuli such as faces (Kreiman et al, 2002; Perrett et al., 1992; Tanaka, 1996).

The primary purpose of these high-level feature detectors is to identify previously seen sets of inputs that correspond to behaviorally relevant objects or people (Barlow, 1995; Page, 2000). This functionality is extremely useful because other information that is relevant to the input pattern can be activated by the appropriate feature detector. This "base level" grouping (Roelfsema, 2006) is achieved during the initial feedforward sweep of activity through the system, during which the visual input ascends through the visual hierarchy activating any relevant feature detectors as it goes. This can enable detection of visual stimuli after only very brief exposure (VanRullen and Thorpe, 2001; Keysers et al., 2001).

However, if the visual scene is noisy or ambiguous in some way, the feedforward sweep may not be sufficient to accurately interpret the scene (Hochstein and Ahissar, 2002; VanRullen et al., 2005; Evans and Treisman, 2005). To aid segmentation, the visual system channels information from higher levels back down the hierarchy, so that activated higher layer feature detectors are able to bias the activity in the lower layers towards the most strongly detected configuration (Vecera and O'Reilly, 1998; Hochstein and Ahissar, 2002).

An obvious weakness of the feedforward hierarchical grouping system is that there are nothing like the number of cells available in the brain as would be required to have a feature detector for every conceivable pattern of input. The solution is, of course, that these cells are not hardwired but actively learn to respond to relevant and useful stimuli. It is this mechanism of learning, and the subsequent top-down affects it has on figure-ground assignment, that is explored in this paper through a behavioral study and a neural network model.

Behavioral Study

A bi-stable figure-ground array was used to demonstrate the effect of familiarity on the process of regional grouping and segmentation. The Gestalt psychologist Rubin first developed this type of stimuli, an early example being his famous faces and vase display (Rubin, 1915/1958), in which it is only possible to see either the faces or the vase at any one time, despite the fact that both are portrayed simultaneously. The region that is consciously perceived is

described as the figure while the remaining region is described as the ground (Rubin 1915/1958). There are two main phenomenological consequences that follow from this figure-ground assignment. Firstly, the figural region appears to have shape, while the ground region appears shapeless (at least near the contour). The second is that the ground region appears to continue behind the figure, giving the impression that the figure is closer.

The Gestalt psychologists identified several fundamental stimulus parameters that would increase the probability that a region would be perceived as the figure. They showed that shapes that were relatively high in contrast, were small compared to the neighboring region(s) or that were enclosed or symmetrical, would be more likely to be seen as the figural region in the image (Rubin, 1915/1958; Koffka, 1935). They believed that a pre-processing stage first grouped and segmented the scene into more manageable units on the basis of these characteristics, and that only following this stage could the influence of experience come into play.

However, contrary to the pre-processing hypothesis, even some of Rubin's early work suggested the existence of a familiarity component in the earliest stages of figure-ground assignment. Rubin noticed that participants that were tested on the same stimuli for a second time would often report seeing the same region as the figure as they had on their previous viewing (Rubin, 1915/1958). He called this phenomena; "a figural after-effect".

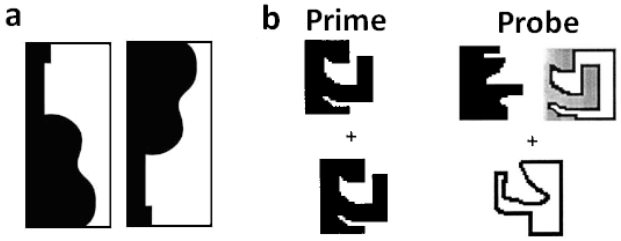

Fig. 1. (a) Peterson & Gibson's denotivity manipulated stimulus. (b) Peterson & Lampignano's stimulus.

More recently Peterson and Gibson (1993) demonstrated the early involvement of familiarity using bi-stable arrays in which only one side of the contour had a familiar or denotative shape. Each image was seen in both its upright and inverted orientations, see figure 1(a). Peterson and Gibson showed that the denotative side of the display was perceived as the figure more often when it was displayed in its upright orientation than when it was displayed inverted. As the low-level Gestalt configural cues such as size and convexity were constant across conditions the perceptual bias must have been generated by the familiarity manipulation.

Furthermore, a priming study conducted by Peterson and Lampignano (2003) has shown that a familiarity bias for a novel shape can be induced after just a single presentation. In their study, the prime display consisted of two novel shapes (boxes with jagged edges on one side) above and below the fixation point, see figure 1(b). The shape above fixation was the prime, and to insure the participant's processed it fully they were required to perform a matching task in which they had to report whether the two shapes were the same or different. Following an inter stimulus interval (ISI) of 1700ms during which time the screen was blank, the probe display would appear. The probe array consisted of two shapes above the fixation point and one shape below. Of the two upper shapes, one was the probe and the other a distractor that was located where the previous prime had been. The task was again to perform a matching judgment, this time based on the white shapes in the array.

In the experimental trials the probe shape was the ground of the prime shape (i.e. had the same contour but the enclosed region was located on the opposite side), while in the control trials the probe shape had a completely different contour to the prime. The authors believed that in the experimental trials the familiarity bias induced by the prime display should induce a preference for interpreting the outside of the probe shape, the shaded area in figure 1(b), as the figure. This influence would be in direct opposition to the interpretation that the traditional Gestalt grouping principles would favor. They hoped that this conflict would slow the participant's ability to resolve the figural region, which would in turn manifest as a slowing in their matching response times. Their experimental findings confirmed this hypothesis.

A problem with this study however, is that in neural terms, the processing advantage found when the current figure is the same as was seen on the immediately preceding trial, may largely be due to the residual activity that is still present in the network. This advantage may be apparent even when little actual learning, implemented via weight changes in the network has taken place.

Despite this issue the evidence for the presence of a familiarity component in early figure-ground processes is compelling. It is difficult to incorporate these findings into a purely feedforward framework such as the Gestalt pre-processing model, but the interactive framework advocated here can easily explain these results. In our model partially segmented information can be passed up the network hierarchy to higher object knowledge regions (feature detectors), which can in-turn feed useful information back to lower level areas where the segmentation is still taking place, allowing object information to bias the process before the final interpretation is reached.

The study presented here is an extension of Peterson and Lampignano's priming study. In the current version the separation between the prior exposure to the novel stimuli (that generates the required familiarity bias) and the subsequent segmentation task is substantially increased. This modification will completely rule out any possibility that the observed effect is due to residual activity in the system, and should therefore constitute a more convincing demonstration of a rapidly acquired familiarity bias.

Method

The paradigm used in this study was a modified version of an experiment conducted by Driver & Baylis (1996), in which they presented the participants with a rectangle that was divided into two regions by a jagged contour with five steps. In their study one of the regions in the rectangle was biased using the Gestalt cues of size and luminance (was smaller and brighter) to be perceived as the figure. They then presented the participants with a pair of probes that appeared either on the same side as the Gestalt favored figural region (figure probes) or on the opposite side (ground probes). The participants had to report which of the two probes had the same contour as in the previous array. They found that participants were faster to match the figure probes and interpreted this result as demonstrating that the region perceived as the figure is remembered while the ground is not; the reason it took longer to find the match in the ground probes they believed, was because it involved a reversal of the encoded figure.

In the current version of the experiment, the same probing technique was used to measure which side had been perceived as the figure, but instead of Gestalt biases, familiarity was used to determine which side of the bi-stable array should be seen as the figure. To induce the familiarity bias (at a longer separation than was used in the study by Peterson and Lampignano) the participants were taught a small set of novel shapes prior to the segmentation phase of the experiment. Then in the segmentation phase the bi-stable array was balanced for Gestalt cues, so the only remaining bias came from the experimental manipulation that the central contour matched one of the previously learnt shapes. The figure probes were located on the same side as the region that matched one of the previously learnt shapes, see figure 2, while the ground probes were located on the opposite side. If the initial learning phase had successfully trained the participants, this familiarity should have biased them to segment the ambiguous arrays in favor of seeing the previously learnt shape as the figure. This should then, as in Driver & Baylis's study, manifest as a processing advantage for the figure probes over the ground probes.

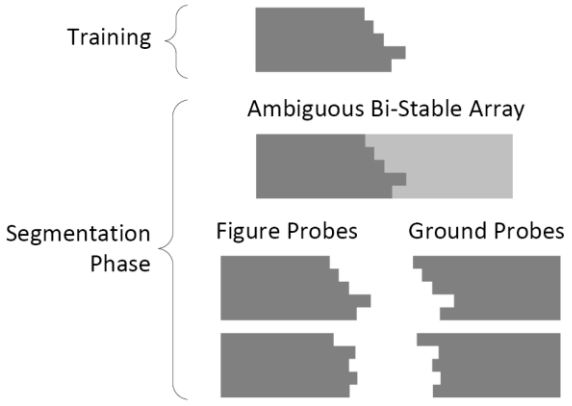

Fig. 2. Behavioral stimuli. The two sides of the bi-stable array were equiluminant, but are depicted here at different contrasts for illustrative purposes.

Participants 26 University of London students aged between 18 and 30 years old participated in the study. All were right handed, and reported normal or corrected to normal visual acuity and color vision.

Stimuli The shapes were balanced for three Gestalt factors: luminance, size and convexity. Luminance was balanced by selecting equiluminant colors using a point light meter. To balance size each shape was set to have exactly the same surface area (equal to half the ambiguous array), and to balance for convexity each participant was taught a separate set of shapes, so that any random convexity biases would balance out across the study.

An additional issue concerning initial fixation was also addressed. Prior to the ambiguous rectangle being displayed the participants were required to focus on a centrally located fixation point, in order to centralize their attention. If either region in the subsequent ambiguous array occupied this central location the participants may have been biased towards perceiving that region as the figure. To remove this possibility the middle step was always located in exactly the center of the rectangle in the same location as the vertical bar of the fixation cross.

With the participants seated 60cm from the screen the ambiguous array was 16.7º wide and 4.3º high. The two probes were either left or right justified and were separated vertically by 0.86 º (The same height as a single step).

Design The design was within-subjects with four factors: Prime color, Prime side, Probe type (figure or ground) and Target probe (top/bottom). Each factor had two levels, resulting in 16 conditions. The training set was comprised of just two shapes, one facing either way, so that the learning task was as easy as possible for the participants. A set of novel shapes were included in the segmentation phase of the experiment to dilute the frequency and repetition of the presentation of the two familiar shapes that might have otherwise given away the motivation of the study. Four distractor shapes were included so the familiar shapes were in the minority.

Procedure The experiment was divided into two sections. The first section was the training phase and the second the segmentation and speeded matching task. The training phase was self-paced, and comprised of two sections. In the first section the participants were shown the two shapes one at a time, in order to first learn them. They then moved onto a multi-choice phase, in which they had to identify the familiar shape from a selection of shapes on the screen. There were two multi-choice presentations, one for each shape, following which the cycle would start again and the participants could re-examine the target shapes. Initially the multi-choice was out of just two shapes but after the participants had correctly identified the target shapes 10 times the number of shapes to pick from was increased to four. The training phase was completed once they had correctly selected the target shapes 20 times from this second multi-choice phase. This training section took between 20 and 30 minutes to complete.

In the segmentation phase of the experiment, each trial started with a centrally located fixation cross (1500ms), that the participants were required to fixate on in order to initially centralize their attention. This was followed by the presentation of the ambiguous bi-stable array (1000ms). The participants were instructed to remember just the contour of the array. The experimenter strictly avoided any reference to the shapes that were formed on either side of the contour. The ambiguous array was then removed and following a brief ISI (500ms) in which the screen was blank except for the fixation cross, the two test probes would appear, one above the other either on the left or the right hand side of the screen (1000ms). The stimuli timings were determined by an earlier pilot study that showed that these values yielded the largest effect. The participant's task was to report, via a button box, which of the two probes had a matching contour to that in the previously seen bi-stable array. They were asked to respond as quickly and as accurately as possible, and if they didn't respond before the probes were removed, the comment 'Too Slow' was displayed in the centre of the screen to encourage them to respond faster in subsequent trials. The segmentation phase of the experiment was made up of 5 blocks each comprised of 48 trials (3 pairs of shapes × 16 conditions).

Treatment of results Initially the data from the first of the five blocks for each participant was discarded as a practice, and the error trials and time-outs were removed from the reaction time analysis. Then two further error reduction procedures were performed. Firstly, participants whose responses did not significantly differ from what could be expected by chance responding alone were removed from the analysis. This was calculated using the binomial probability distribution. The maximum number of errors that were allowable before responding did not significantly differ from chance levels with 64 trials was calculated to be 24, or 37.5% errors (as $C(64,24)\times0.5^{24}\times0.5^{40}=0.029$, but $C(64,25)\times0.5^{25}\times0.5^{39}=0.052$). Secondly, data points that were beyond two standard deviations from a participant's median response time (calculated recursively), were removed. This process removed excessively early or late response's (that may signal false positive responses). Six participants made more than the threshold number of allowable errors and so were removed from the data set. A further four had missing data for one or more condition(s) following the removal of their outlying data points. This left 16 participants in the final analysis. The removal of outlying data points excluded 8.79% of the final 16 participants recorded data. Four 4-way ANOVAs were run, one on the RT data and one on the accuracy data, for both the raw and the processed data sets. The results were qualitatively identical, so only the findings from the processed data set will be described any further.

Results

The means of participant's median reaction times and corresponding accuracy data for figure- and ground-probes can be seen in figure 3.

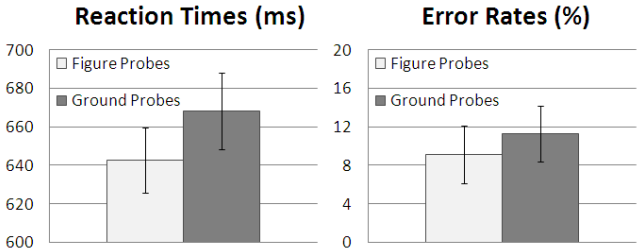

Fig. 3. RTs and percentage errors in each condition. Error bars show within-subjects 95% confidence intervals.

The ANOVA on the RT data found the expected main effect of probe type ($F(1,15) = 9.870$, $p = .007$), with figure probes being processed 25.50ms faster than ground probes (figure probes = 642.88, ground probes = 668.38).

The error ANOVA however, did not find a significant difference between the two conditions ($F(1,15) = 0.427$, $p = .523$), although the difference was in the expected direction with 2.15% less errors made for figure probes (figure probes = 9.18%, ground probes = 11.33%).

The RT result demonstrates a rapidly acquired familiarity bias for the novel shapes that can only be the result of actual learning as opposed to priming.

Modeling Work

To simulate the behavioral work a simple two layered network was built. The nodes in the lower layer were responsible for representing the inputs (in a one-to-one fashion) while the nodes in the upper layer were responsible for learning the re-occurring patterns of activity (shapes) in the lower layer.

The behavioral stimuli were represented using a 6 by 5 grid. A unit high for each step and wide enough to allow sufficient variability between shapes. Edge inputs were used instead of the figure units themselves, in order to faithfully replicate the contour detecting cells found in V1 (Field et al., 1993; Zhaoping, 2003).

The 6 by 5 input grid contained 71 edge elements, so the lower layer of the neural network contained 71 nodes, one to represent each edge, see figure 4(b). The network was required to learn two shapes as in the behavioral work, and since each pattern can be encoded by a single node, just two upper layer nodes were included. The lower layer nodes were then connected in an all-to-all fashion to each of the upper layer nodes. In the cortical hierarchy this convergence would take several more layers to be achieved, but this simplified representation was used as it minimizes the networks complexity whilst still capturing the core principles involved.

The network was then trained by repeatedly presenting the shapes in a random sequence to the network inputs. The patterns were learnt by adjusting the feedforward weights of the network using an activity dependent learning rule that strengthens the weights between sets of co-active lower layer and upper layer nodes (see Spratling & Johnson, 2004 for details). So that only a single upper layer node would

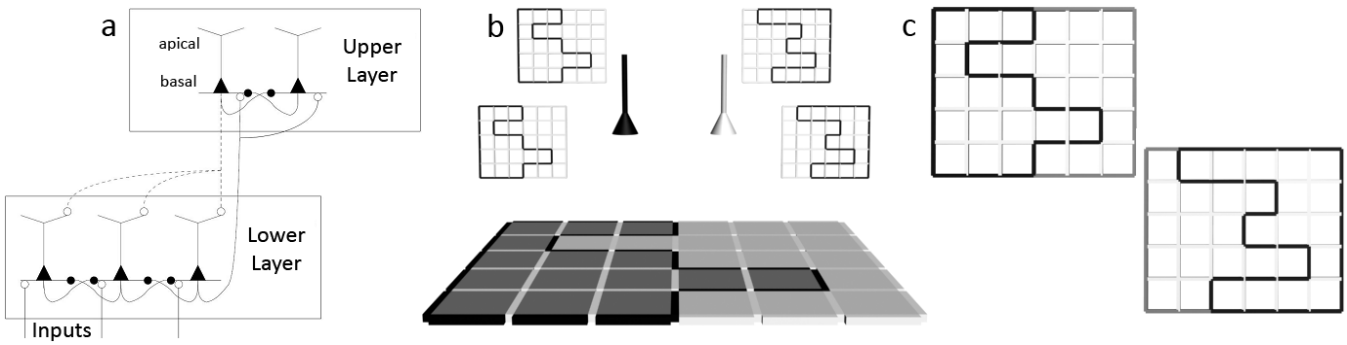

Fig. 4. Neural network model. (a) The connectivity used in the network: Solid lines represent FF connections; Dashed lines represent FB connections; Open circles represent excitatory synapses and Filled circles represent inhibitory synapses. (b) The Network during training. The nodes in the lower layer are not explicitly drawn, but each edge is represented by its own node, and that nodes activity level is represented by the shading of its corresponding edge element, with darker shades representing higher firing rates. The grids next to the upper layer nodes show both the FF (top) and FB (bottom) weight matrices that have been learnt by the network. (c) Activity in the lower layer during testing with the ambiguous bi-stable arrays.

respond to a particular pattern in the lower layer, and therefore would later signal that pattern's presence by its activation, the upper layer nodes were made to compete. The form of competition used in this model was pre-integration lateral inhibition (see Spratling & Johnson, 2002, 2003, for details). At the start of the training process all the weights in the network were set to be equal. This meant that when the patterns were presented to the network for the first time, each upper layer node would become equally active. To resolve this deadlock and ensure a single node would win the competition and therefore go on to represent that input in the future, noise was applied to the output values of the upper layer nodes.

The network described thus far can successfully learn and subsequently recognize the required shapes, but the functionality of specific interest to this study only begins to emerge when the recognition signals that result from this initial feedforward sweep of activity are channeled back down the processing hierarchy so they can in turn affect the lower level representations of the scene.

A central assumption of this work is that conscious perception is generated by long range recurrent activity that connects frontal regions with the most salient stimuli in the earliest layers of the visual cortex, where spatial resolution is at a maximum and the visual scene is represented most clearly. All neurons along the recurrently activated pathway experience enhanced activities, and as a result these enhancements can be considered as direct physiological markers of perception (Lamme, 2007; Roelfsema, 2006; Lamme & Roelfsema, 2000).

In order for the activation of a higher layer feature detecting node to preferentially enhance its favored input pattern (so that it is more likely to become the target of the longer range recurrent activity that underlies perception), it must develop reciprocal feedforward and feedback weights. This is achieved using a similar activity dependent learning rule for the feedback weights as was used for the feedforward weights, simply with the pre- and post-synaptic values reversed. The resulting feedback signals can then modulate the activities of the nodes in the lower levels of the network.

The form of modulation used in this model is multiplicative as it allows the feedback signals to enhance bottom-up activity that already exists, whilst preventing the generation of illusory activity that can result from additive techniques. This results in a system that is strongly stimulus driven, but that can be biased/focused by the influences of top-down effects. The model is constructed using nodes with two integraton sites, rather than the point-neuron models used in many other simulations. This more accurate model of the pyramidal cell's physiology allows the separate feedforward and feedback signals to be processed independently before they are finally combined to determine the output of the neuron (Kording & Koenig, 2000; Kording & Koenig, 2001; Larkum & Sakmann, 1999). Inspired by observed anatomy, the feedforward connections target the basal dendrites, while the feedback connections target the apical dendrites (Budd, 1998; Rockland, 1998; Rolls and Treves, 1998), see figure 4(a) for a network schematic.

A different integration function is required to calculate the activation of the apical dendrites to that used to calculate the activation of the basal dendrites. When information is travelling up the visual hierarchy, the node's task is to group the incoming inputs; a function that is best performed using the weighted sum. However, when information is traveling back down the processing hierarchy, the inputs to each lower layer node will originate from multiple feature detectors that represent discrete stimuli. Co-active feedback inputs should therefore not be pooled; rather the apical dendrite's value should reflect the most active upper layer feature detector, as this will represent the most likely interpretation. An appropriate activation function for the apical dendrites should therefore return the maximum of the pre-synaptic input values. For full mathematical details of the model's implementation please refer to: Spratling, M.W. & Johnson, M.H. (2004).

Once the network had been trained it was tested using stimuli consistent with the behavioral experiment's bi-stable

array, comprised of one of the training shapes embedded within a full outlined border. To evaluate the models perceptions, it was not necessary for it to perform the matching task used in the behavioral experiment, rather the more direct, physiological correlate of figural assignment, enhanced regional activities, could be used.

Results

The network successfully learnt the required reciprocal feedforward and feedback weight matrices, see figure 4(b). Then when the network was presented with an ambiguous array, the upper layer node that corresponded to the embedded familiar shape became more strongly activated than the remaining node, which was only partially activated by the border elements on the non-familiar side of the array that were part of its favored shape. This differential activity was then sent back to the lower layer where the familiar shape received preferential enhancement and hence could be considered to have been perceived as the figure. Figure 4(c) portrays the activities of the lower layer nodes following the presentation of each ambiguous array. The nodes activities are represented in grayscale with darker shades representing higher activities.

Conclusion

The behavioral component of this study demonstrated the rapidity with which object representations can be acquired, and the biasing effect they can subsequently exert on the process of figure-ground assignment. A biologically plausible neural network model was then built, that successfully simulated the behavioral task, and in doing so identified several key computational principles that are in operation in the visual system.

Acknowledgments

This work was funded by an EPSRC grant: number GR/S81339/01.

References

Barlow, H. B. (1995). The neuron doctrine in perception. In Gazzaniga, M. S., editor, The Cognitive Neurosciences , chapter 26. MIT Press, Cambridge, MA.

Budd, J. (1998) Extrastriate feedback to primary visual cortex in primates: a quantitative analysis of connectivity. Proceedings. *Biological sciences / The Royal Society, 265(1400),* 1037-44.

Driver, J. & Baylis, G. (1996) Edge-Assignment and Figure-Ground Segmentation in Short-Term Visual Matching. *Cognitive Psychology, 31(3),* 248-306.

Evans, K. & Treisman, A. (2005) Perception of objects in natural scenes : Is it really attention free? *Journal of Experimental Psychology: Human Perception and Performance, 31*,1476--92.

Field, D.J. Hayes, A. & Hess, R.F. (1993). Contour integration by the human visual system: Evidence for a local association field. *Vision Research, 33*, 173-193.

Gilbert, C.D. Das, A. Ito, M. Kapadia, M. & Westheimer, G. (1996) Spatial integration and cortical dynamics. *PNAS, 93*, 615-622.

Hochstein, S. & Ahissar. M. (2002) View from the top: hierarchies and reverse hierarchies in the visual system. Neuron, 36(5):791--804.

Kapadia, M.K. Ito, M. Gilbert, C.D. & Westheimer, G. (1995). Improvement in visual sensitivity by changes in local context: Parallel studies in human observers and in V1 of alert monkeys. *Neuron, 15*, 843-856.

Koffka, K. (1935). *Principles of Gestalt psychology*. New York: Harcourt Brace.

Körding, K. & König, P. (2000) Learning with two sites of synaptic integration. *Network: Comput. Neural Syst., 11(1),* 25-39.

Körding, K. & König, P. (2001) Supervised and Unsupervised Learning with Two Sites of Synaptic Integration. *Journal of Computational Neuroscience, 11(3),* 207-215.

Kreiman, G. Fried, I. & Koch, C. (2002). Single-neuron correlates of subjective vision in the human medial temporal lobe. *PNAS, 99*, 8378-8383.

Lamme, V.A.F. (2006) Towards a true neural stance on consciousness. *Trends in Cognitive Science, 10(11):*494-501. Review.

Lamme, V.A.F. & Roelfsema, P.R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neuroscience, 23*, 571-579.

Larkum, M. & Sakmann, B. (1999) A new cellular mechanism for coupling inputs arriving at different cortical layers. *Nature, 398(6725),* 338-41.

Page, M. (2000). Connectionist modelling in psychology: a localist manifesto. Behavioural and Brain Sciences, 23(4):443–67.

Perrett, D. I., Hietanen, J. K., Oram, M. W., & Benson, P. J. (1992). Organisation and functions of cells responsive to faces in the temporal cortex. Philosophical Transactions of the Royal Society of London , 335:23–30.

Qiu, F.T. & von der Heydt, R. (2005) Figure and Ground in the Visual Cortex: V2 Combines Stereoscopic Cues with Gestalt Rules. *Neuron, 47*, 155-166.

Rockland, K. (1998) Complex microstructures of sensory cortical connections. *Current opinion in neurobiology, 8(4),* 545-51.

Roelfsema, P. R. (2006) Cortical algorithms for perceptual grouping. *Annual Review of Neuroscience, 29*,203--27.

Rolls, R.T. & Treves, A. (1998) Neural Networks and Brain Function *Oxford University Press*.

Rubin, E. (1958). Figure and ground. In D. Beardslee & M. Wertheimer (Ed. And Trans.), *Readings in perception.* (pp. 35-101). Princeton, NJ: Van Nostrand. (Original work published in 1915).

Spratling, M.W. & Johnson, M.H. (2002) Pre-integration lateral inhibition enhances unsupervised learning. *Neural Computation,, 14(9),* 2157-79.

Spratling, M.W. & Johnson, M.H. (2003) Exploring the functional significance of dendritic inhibition in cortical pyramidal cells. *Neurocomputing ,52-54*, 389-95.

Spratling, M.W. & Johnson, M.H. (2004) A Feedback Model of Visual Attention J. *Cognitive Neuroscience, 16*, 219-237.

Tanaka, K. (1996). Representation of visual feature objects in the inferotemporal cortex. Neural Networks , 9(8):1459–75.

Vecera, S.P. & O'Reilly, R. (1998) Figure-ground organization and object recognition processes: An interactive account. *Journal of Experimental Psychology: Human Perception and Performance, 24*,441--62.

VanRullen, R. & Thorpe, S. J. (2001) Is it a bird? is it a plane? ultra-rapid visual categorisation of natural and artifactual objects. Perception, 30:655--68.

Keysers, C. Xiao, D. K. Földiák, P. & Perrett, D. I. (2001) The speed of sight. Journal of Cognitive Neuroscience, 13(1):90--101.

VanRullen, R. Reddy, L. & Fei-Fei. L. (2005) Binding is a local problem for natural objects and scenes. Vision Research, 45(25-26):3133--44.

Peterson M. A. & Gibson B. S. (1993). Shape recognition inputs to figure-ground organisation in three-dimensional displays. Cognitive Psychology, 25:383—429.

Peterson M. A. & Lampignano, D.W.(2003). Implicit memory for novel figure-ground displays includes a history of cross-border competition. *Experimental Psychology: human Perception and Performance, 29*, 808-822.

Zhaoping, L. (2003) V1 mechanisms and some figure-ground and border effects. *Journal of Physiology-Paris, Neuroscience and Computation, 97*, 503-515.