



PERGAMON

Vision Research 38 (1998) 3095–3107

**Vision
Research**

Learning pop-out detection: building representations for conflicting target-distractor relationships

Merav Ahissar^{a,*}, Roni Laiwand^b, Gal Kozminsky^b, Shaul Hochstein^{b,c}^a Department of Psychology, Hebrew University, Mt. Scopus Campus, Jerusalem 91905, Israel,^b Neurobiology Department, Institute of Life Sciences, Hebrew University, Givat Ram, Jerusalem 91904, Israel^c Center for Neural Computation, Hebrew University, Givat Ram, Jerusalem 91904, Israel

Received 28 June 1996; received in revised form 25 November 1996; accepted 10 November 1997

Abstract

Studies of perceptual learning consistently found that improvement is stimulus specific. These findings were interpreted as indicating an early cortical learning site. In line with this interpretation, we consider two alternative hypotheses: the ‘earliest modification’ and the ‘output-level modification’ assumptions, which respectively assume that learning occurs within the earliest representation which is selective for the trained stimuli, or at cortical levels receiving its output. We studied performance in a pop-out task using light bar distractor elements of one orientation, and a target element rotated by 30° (or 90°). We tested the alternative hypotheses by examining pop-out learning through an initial training phase, a subsequent learning stage with swapped target and distractor orientations, and a final re-test with the originally trained stimuli. We found learning does not transfer across orientation swapping. However, following training with swapped orientations, a similar performance level is reached as with original orientations. That is, learning neither facilitates nor interferes to a substantial degree with subsequent performance with altered stimuli. Furthermore, this re-training does not hamper performance with the originally trained stimuli. If training changed the earliest orientation selective representation (specializing it for performance of the particular performed task) it would necessarily affect performance with swapped orientations, as well. The co-existence of similar asymptotes for apparently conflicting stimulus sets refutes the ‘earliest modification’ hypothesis, supporting the alternative ‘output level modification’ hypothesis. We conclude that secondary cortical processing levels use outputs from the earliest orientation representation to compute higher order structures, promoting and improving successful task performance. © 1998 Elsevier Science Ltd. All rights reserved.

Keywords: Perception; Vision; Detection; Search; Training; Learning; Modification; Behavior; Psychophysics

1. Introduction

The dramatic effects of practice in improving the performance of adult humans in even the most simple perceptual tasks have been studied recently using several paradigms [1–17]. The consistent finding was that learning effects are substantially specific to the trained parameters along such basic visual dimensions as orientation, size (or spatial frequency) and retinal position (although there is substantial inter-subject variability). The studies concurred in concluding that learning

occurs at a level that maintains separate representations along these dimensions. If one assumes that electrophysiological data describing single unit receptive field selectivities reflect stimulus representations at the various cortical areas [18], these results imply that learning occurs at an early cortical area, perhaps V1 or V2.

None of these studies, however, attempted to study directly the mechanisms underlying the measured improvement. In the present study, we examine the feasibility of the following simple hypothesis, which is consistent with the findings of stimulus specificities:

Learning occurs within the first neuronal level whose representation is discriminative with respect to the required task.

* Corresponding author. Fax: +972 2 5851159; e-mail: msmerava@pluto.huji.ac.il.

Earliest modification has been implicitly assumed in models of perceptual learning for various simple visual tasks [19,20,12]. We now show that the pattern of the improved behavior could not be obtained without modifications in the output levels, disproving any ‘hard version’ of the ‘earliest modification’ hypothesis. We thus suggest an ‘output level modification’ hypothesis, namely:

Learning occurs within neuronal levels that read the output of the first neuronal level whose representation is discriminative with respect to the required task.

A more specific account of which output levels are modified under different training conditions is described elsewhere [21]. Briefly, we suggest that learning proceeds in the reverse direction of the visual processing hierarchy. Rather than occurring at the earliest discriminative level, learning occurs at the highest level that retains discriminability with respect to the required visual dimension, at a default degree of analysis precision. Lower level representations will only be modified (as part of the learning process) when higher levels are not sufficiently discriminative, and if training is performed so that the appropriate low-level population is accessible.

For the purposes of these ‘earliest modification’ and ‘output level’ hypotheses, we infer cortical representation properties from published single neuron receptive field characteristics. For example, consider a perceptual task requiring discrimination between vertical and horizontal light bars, where each type of bar is presented in random position anywhere within the visual field. We would conclude that this task is not accomplished within a retinal representation, since retinal neurons are not differentially activated by these two types of stimuli (vertical and horizontal). Rather, the task must be achieved at neural levels that explicitly code for orientation. Looking for a cortical site, however, one finds a succession of cortical areas with representations which differentiate stimulus orientations (at least all the areas from primary area V1 to the infero-temporal area, IT). The ‘earliest modification’ assumption asserts that learning orientation-based discriminations occurs at the first level that explicitly codes for orientation, namely the primary visual cortical region V1 [22]. Similarly, the ‘output level modification’ hypothesis would predict that learning orientation-based tasks occurs at one or more of the levels receiving output from this early V1 level, including ‘higher’ V1 levels, V2, V3, V4 and MT.

We test the earliest modification assumption in the context of early perceptual learning for the case of detection of an oddly oriented target element in an array of homogeneously oriented distractor light bars. When the target greatly deviates in its orientation from

that of the distractors, it ‘pops-out’ so that its detection is simple and pre-attentive [23,24]. Physiological data indicate that the responses of a large proportion of the neurons, at the level of area V1 and above, are tuned around their preferred or ‘best’ orientation (V1: Hubel and Wiesel, [22]; review for extra-striate areas: Desimone and Ungerleider, [18]). Indeed there is anatomical [25], physiological [26] and behavioral evidence (e.g. locality: Nothdurft, [27]; Sagi and Julesz, [28]) suggesting that pop-out detection may be accomplished within area V1.

In previous studies we found that pop-out detection is learnable: its performance substantially improves with practice [10,11,29–32]. The ‘earliest modification’ hypothesis asserts that learning derives from modifications of interactions between neurons at the earliest stage which explicitly codes for orientation. Thus, according to this concept, learning in this case is the gradual modification of interactions among the orientation selective neurons that are tuned around the trained (target and distractor) orientations. Our previous findings, that the improvement is largely specific to the orientations of the target and distractor elements, to the location of the target, and to element size [31,11] are compatible with improvement deriving from specific modifications of lateral interactions between local, orientation and size selective neurons.

A simple mechanism implementing modifications of lateral interactions would be of increasing lateral inhibition between the distractor detectors. This would decrease the overall response of distractor detectors and would consequently increase target salience. We found, however, that learning is highly specific also to target orientation [11] refuting this possibility. Thus, while the simplicity of pop-out detection may result from a ‘winner takes all’ mechanism [33] based on lateral inhibition between neurons with similar orientation selectivity [26,34,35], its improvement as a function of practice does not stem from modifications within these connections. Given that modification of interactions between distractor detectors alone was refuted, we examine, in this study, an alternative option for implementing the earliest modification hypothesis. We test for modifications of interactions between target and distractor detectors, as subserving pop-out.

The procedure we use to test the earliest modification hypothesis is to examine ‘second’ learning, i.e. learning with modified stimulus parameters. The stimulus modification we use is swapped target and distractor orientations, a consistently strong interfering modification. We measure performance and learning with the swapped orientations (following training with the original orientations), as well as the effect of the second learning on subsequent performance with the original stimuli. The logic underlying this procedure is further explained below (see Rationale). The major result is

that, although subjects do not transfer learning across orientation swapping, they are not confused by it. Furthermore, subsequent training leads subjects towards a similar asymptote, and yet does not disrupt performance with the originally trained stimuli. That is, learning with swapped orientations does not hamper the original training achievements.

We argue that the lack of asymptotic interference invalidates the ‘earliest modification’ assumption. Changes in early V1 orientation tuning properties and/or intra-area neuronal interactions, which favor and improve detection for one set of parameters, would necessarily have an interfering effect on performance of a task using swapped orientations, unless compensated by substantial modifications of the read-out of this early representation. As no such interference was found, learning must affect output connections of an orientation selective representation. For example, learning may affect an explicit representation of orientation difference between target and distractor orientations, occurring at a higher cortical level [36]. Accordingly, we propose the alternative ‘output-level modification’ hypothesis.

2. Methods

2.1. Stimuli and procedure

Stimuli were arrays of light bar elements (147 cd/m²) on a dark background (0.2 cd/m²). The array consisted of 7 × 7 elements (subtending 4.6 × 4.6°) centered around fixation, as illustrated schematically in Fig. 1, top. Each stimulus element subtended 22 × 1’. The distance between element centers was 42.6’ (± 4’ jitter, randomly chosen with uniform probability). In half of the stimulus presentations, all elements had the same orientation. In the other half, one of the elements was a target at a fixed orientation, deviating by 30 or 90° from that of the distractor elements. The orientations used were as follows: For 30° difference: distractors: 30° (or 60°) counter-clockwise from horizontal; target 60° (or 30°); For 90° difference: distractors: 30° (or 120°); target 120° (or 30°).

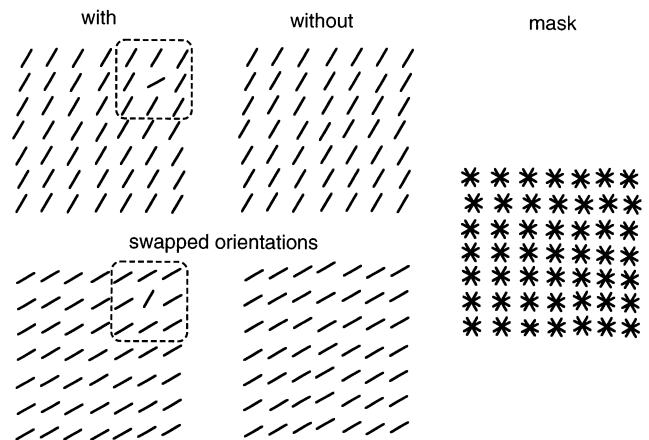
A mask followed each stimulus, as shown in Fig. 1, top, right. The mask was composed of a 7 × 7 array of asterisk-like elements, located at the grid points of the 7 × 7 stimulus lattice (± 4’ jitter so that element position exactly matched those of the stimulus). Each mask element was a superposition of 4 lines: the trained target and distractor orientations, and these orientations plus 90° (e.g. 30, 60, 120 and 150°).

The temporal sequence of each trial is shown in Fig. 1, bottom: Each trial started with a fixation cross (a + sign with 22 × 1’ lines of intensity 147 cd/m²). When the observer pressed the ready key, following

120–165 ms, the stimulus appeared. After a variable duration (the stimulus onset asynchrony, SOA), the mask was displayed for 166 ms. Stimulus presentation was on until mask presentation in the main subject group (30° difference), and lasted 16 ms in the paradigm of 90° difference (where a dark period separated between stimulus offset and mask onset, as shown in Fig. 1, bottom). Finally, following a 233 ms dark period, the fixation point reappeared while the subject pressed a response key. A computer tone confirmed correct responses.

Stimuli were presented in blocks of 20 trials with the same SOA. Each session comprised 70 blocks (1400 trials). Each session began with a set of 9 blocks starting from the longest SOA (183 or 150 ms) and gradually reaching the shortest SOA (16 ms) in an interleaved manner (blocks with SOA of 183, 133, 100, 66, 33 ms, or blocks of 16 ms shorter times, followed by blocks of 150, 116, 83, and 50 ms, or blocks of 16 ms

stimuli for pop-out detection



trial temporal sequence

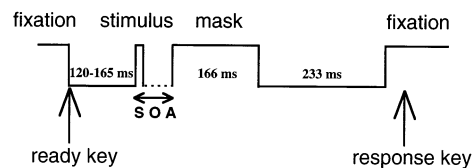


Fig. 1. Schematic diagram of pop-out testing paradigm. Top: A 7 × 7 array of light bars was presented briefly, either with (left), or without (center), an odd element, i.e. a bar with a different orientation. A mask comprising bars of various orientations (right) was presented following each stimulus. Following training with the original set of element orientations, subjects were tested and trained with stimuli in which the target and distractor orientations were swapped (second row). Pop-out target element could appear in any location in the 7 × 7 matrix except the central (fixation) position. Bottom: Trial temporal sequence; note brief stimulus and variable Stimulus Onset Asynchrony (SOA) between stimulus and mask. See text for details. The stimulus areas demarcated by the dashed lines are reflected in the cortical set of neighboring hypercolumns of the schematic diagrams of Figs. 2 and 3.

shorter times). Based on performance in these initial blocks, the range of SOAs to be presented next was chosen so that the shortest SOA would be the longest in which the subject still performed near chance level (55% correct) and the longest SOA would be the shortest where the subject already showed near perfect performance (95% correct). Within that chosen range (constrained to include at least three different SOAs), blocks were presented in pseudo-random sequence. Following blocks of presentations with these SOAs, the next range of SOAs was chosen based on performance in these blocks and following the above criteria. As a result of this procedure, performance was kept around 75% correct, within and throughout sessions.

Stimuli were presented on an HG Trinitron Multi-scan monitor (Sony) or a 5A Micro-scan monitor (A.D.I.) running at 60 Hz frame-rate and 1024×1024 pixel resolution, driven by a #9-GX graphics card (#9 Computer) in a 486 PC computer. Response keys were '1' (for present) and '0' (for absent) on the numeric keypad of the computer keyboard, followed by the ready key, 'enter', to initiate the next trial.

2.2. Subjects

Thirty two subjects participated in the experiments. Subjects were 20–27 years old, with normal or corrected to normal eye-sight. All were naive to the purposes of the experiment and were paid for participation.

2.3. Analysis

The average session threshold was evaluated by computing the best fit psychometric function of the form: $f = 1 - 0.5 \exp(-t/\tau)^{\sigma}$ where f is the fraction of correct responses; t , the trial SOA; and τ and σ are free parameters: τ the threshold SOA at 81.6% correct, and σ , the slope at threshold multiplied by $2e$ [37]. Determination of the threshold is described elsewhere [10,11].

The average performance at each SOA as a function of trial number was computed by summing across subjects the fraction of correct trials within the total number of trials, lumping together the data for all subjects. We also looked at the results averaging the fractions correct for each subject, and found no substantial differences for the two methods of averaging.

The average two-dimensional spatial distribution of fraction of correct detection was computed by summing (across subjects), separately for each position, the number of target-present answers among target-present trials. This summation was performed separately for each SOA. The average was then obtained by simple averaging across a group of SOAs (33, 50 and 66 ms). Thus, performance at each SOA was given the same weight, although the number of presentations was typically not

equal (see Stimuli and Procedure, above and in Ahissar and Hochstein, [11]).

3. Rationale

As outlined in the Section 1, the basic prediction of the 'earliest modification' hypothesis, (and any learning mechanism of gradual modification of interactions among orientation selective neurons tuned around the target and distractor orientations), is that increasing the strength of one type of lateral interaction will reduce the effectiveness of the equivalent lateral interaction in the reverse direction. The reduction could occur in one of two ways: either increasing the strength of one set of connections directly decreases the strength of 'competing' connections (replacement), or increasing the strength of one set of connections leaves the other intact, but the two sets have 'competing' effects (superposition). This interference is schematically illustrated in Fig. 2 for excitatory interactions. Each circle denotes a neuron, or column of neurons, with receptive field(s) tuned around a preferred orientation. The 'best' orientation is indicated by the orientation of the line drawn inside the circle. Arrows indicate unidirectional excitatory connections. During initial learning, the excitatory connections from distractor detectors to target detectors are strengthened, increasing the facilitatory effect of distractor detectors on target detectors, increasing target salience and improving performance (top row; the increase in magnitude of the response is indicated by the increased line width). If, subsequently, performance is tested with a stimulus with swapped target and distractor orientations, there should be no gain from the prior practice; on the contrary, the lateral connections strengthened by previous training would increase the salience of the now distractor detectors, and thus, some initial confusion would be expected (middle row, left)¹.

Second training with swapped orientations, could result either in superimposing reciprocal connections in the reverse direction (left option, Fig. 2, middle row, middle column), or, alternatively, obliterate previously formed connections, replacing them with connections in the reverse direction (right option, Fig. 2, middle row, right column). According to the first scenario, second learning should not approach the same asymptote as the original learning, since its effectiveness is reduced by previously strengthened connections in the reverse direction (strengthening current distractor salience; compare middle row middle column, with top row, right). Furthermore, second learning should disturb asymptotic performance with the original orientations,

¹ We assume that the strategy has not be modified. That is, detection is still by local gradient.

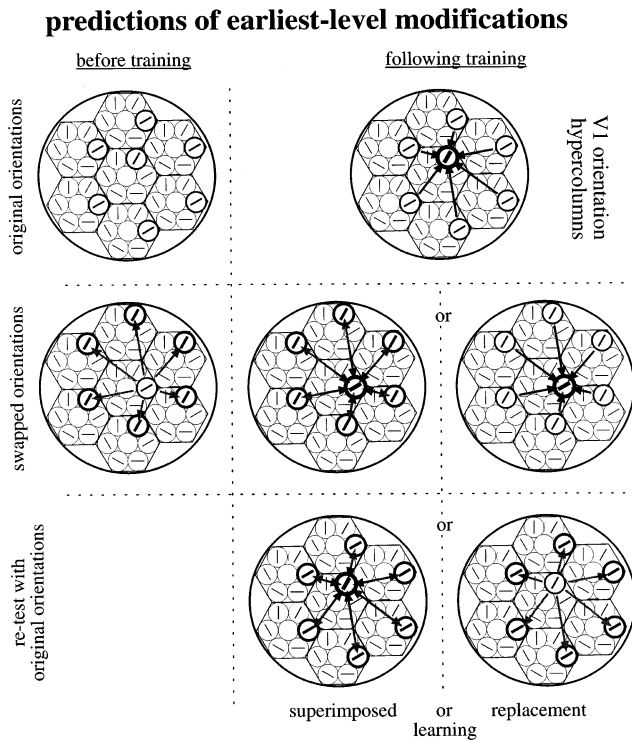


Fig. 2. Schematic illustration of expected effect of lateral interaction model of early perceptual learning according to the ‘earliest modification’ hypothesis. Each hexagon schematically represents a hypercolumn. The orientation preference of each cortical column (circles) is represented by the orientation of the bar drawn within it; the thickness of the bar and circle represents the degree of activation of the neurons in the column. Training enhances specific facilitatory interactions improving performance on this stimulation set. This is demonstrated in the figure by facilitation of target salience by excitatory connections from surrounding distractor detectors (arrows). Training is expected to interfere with subsequent performance of alternative stimulation configurations, especially when the orientations of the target and the distractor elements are swapped. This is seen in the second row (left) where the previously strengthened excitatory connections now enhance the distractor detectors. Second training—with the swapped orientations—would introduce excitatory connections in the reverse direction, either superimposed on (middle column) or replacing (right column) the originally formed connections. In the former case, there would be reduced performance even at asymptote. In this case, the same hampered asymptotic performance will be found for re-testing subjects with the original orientations (compare bottom row, left with top right and middle, middle). In the case of replacement learning, asymptotic performance is always the same, but performance on the re-test is as poor as with swapped orientations before second learning (compare bottom row, right with top right and middle left). See text for further details.

so that a re-test should show a similar performance level as with swapped orientations (Fig. 2, bottom, left)². The other alternative, of replacing connections,

² Here too, the discussion is qualitative. The magnitude of the interference effect is hard to predict. We assume that if there is a large original learning (resulting from strengthening excitatory distractor to target interactions), its interference with asymptotic performance after swapping will not be undetectable. That is, we assume there are no large non-linearities in the additive effects of distractor to target interactions.

predicts that second learning should approach a similar asymptote as the original one, but, would greatly hamper performance with original stimuli, so that performance level on a re-test (Fig. 2, bottom, right) should be worse than naive performance (top, left). In any event, we expect interference between learning swapped and original orientations. Indeed interference occurs in a neural network model implementing pop-out learning by modifications in lateral interactions [20].

Although in Fig. 2 we referred specifically to excitatory connections from distractor to target detectors, an alternative model of increasing inhibition from target to (neighboring, or spatially overlapping) distractor detectors, will have the same effect³. In fact, a similar prediction would be the outcome of any type of asymmetrical interactions between target and distractor de-

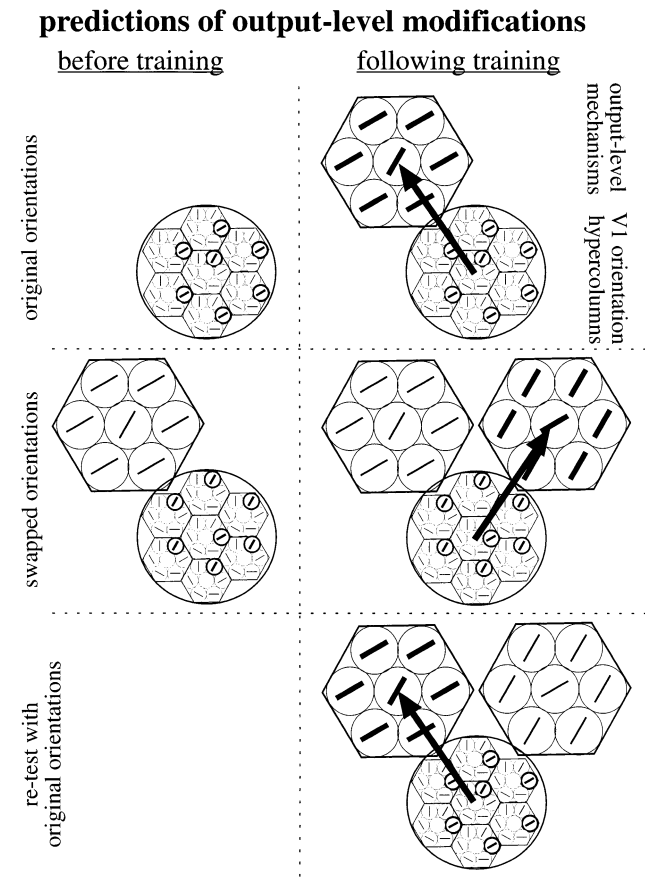


Fig. 3. Schematic illustration of expected effect of learning for output-level modification hypothesis model. Output connection refinements which occur when learning original and swapped orientations develop in different, parallel, locations within the output level and thus do not necessarily interfere with each other, nor affect the earliest orientation representation. See text for details.

³ In fact, in this case, swapping will have a larger effect, since originally one target inhibited several neighboring distractors. With swapping there are many distractors (previous targets) which will inhibit the present target.

tectors: second learning with swapped orientations will always interfere with (be affected by, and/or affect) initial learning.

The alternative ‘output level modification’ hypothesis is described schematically in Fig. 3. Here, training changes connections between lower and higher levels (e.g. connections from V1 to V2 or between levels within one of these areas). The new connections that are appropriate for the original task need not interfere with performance of the task with swapped orientations, since the latter may depend on a parallel mechanism. Similarly, changes in this parallel mechanism will not affect asymptotic performance with the original orientations, on a re-test. To illustrate this hypothesis, we use the same conventions as in Fig. 2. The arrow denotes (refinement and) activation of an output level mechanism. Improved performance following training (top, right) derives from refined selection of appropriate output connections to (or within) a higher level. Stimuli with swapped target and distractor orientations will not activate this mechanism (middle row, left) and second learning will refine a parallel output mechanism (middle row, right). This mechanism need not be hampered by the previously trained mechanism, so that asymptotic performance with the swapped orientations will be similar to that with the original stimuli. Similarly, re-test with the original orientations (bottom) is unaffected by the intervening second learning. The essential characteristic is that changes occur in separate parallel mechanisms, that do not necessarily affect each other, or the earliest orientation representation.

Thus, there are different qualitative predictions from the two alternative hypotheses: interference or independence when training first with one set of orientations and then with swapped orientations and re-testing with the original set. These are examined in the following experiments.

4. Results

4.1. Asymptotic levels with swapped and original orientations

Ten subjects were trained until reaching asymptotic performance levels. Subsequently they were tested and re-trained with swapped orientations. Following several training sessions with the swapped orientations (seven or more), they were given a re-test session with the originally trained orientations. Fig. 4 plots data for three individual subjects and the average for all ten subjects. There was a major performance improvement during the first session with the original orientations, and subsequent practice yielded further improvement [11]. The test with swapped orientations found thresholds nearly at pre-training levels. That is, almost

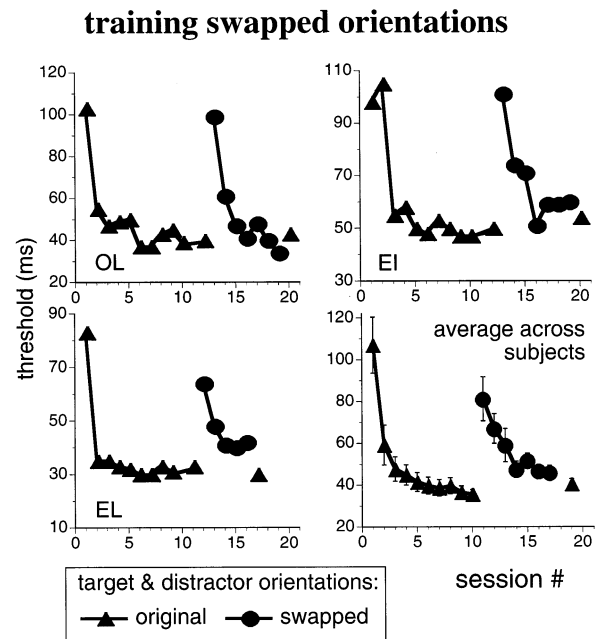


Fig. 4. Learning and interference with swapped orientations. Examples of three subjects and the average for ten subjects (bottom, right) who trained with original orientations (triangles), and were subsequently tested and retrained with swapped target and distractor orientations (circles). Following several training sessions, they were re-tested with the originally trained orientations, to examine whether there was interference (final triangle). Note there is nearly no transfer to the swapped orientations condition (first circle), i.e. learning had to begin again. Furthermore, there is some slowing down of the learning for the swapped condition compared with the original learning. On the other hand, following training with the swapped orientations, there was no interference when re-testing with the original orientation condition—final triangle has just as low a threshold as asymptote before swapped orientation training. This is contrary to the expectations of the earliest modification (lateral interactions) model illustrated in Fig. 2, supporting the alternative output level modification hypothesis of Fig. 3.

no transfer was found to the swapped orientation condition. In fact, for every one of the ten subjects the swapped threshold was substantially above that of the second session with the original orientation, although for most (7/10) it was somewhat below that of the first session (Fig. 7). This specificity is predicted by both hypotheses. However, as pointed out above, any asymmetric interaction would induce also some confusion in the first few trials, following swapping.

We therefore focused further on the first 40 trials with original and swapped orientations. During the first 40 trials, two blocks of long SOAs (> 100 ms) were presented. Fig. 5 illustrates the average percent correct with original (triangles) and with swapped (circles) orientations. While initially, following the first few trials, performance was close to chance level, it started already near optimal for swapped orientations. Not only was there no confusion due to swapping orientations, but performance at long SOAs was immediately and almost fully transferred.

The initial transfer at long SOAs could derive from initial learning of non-orientation specific aspects which are shared by both original and swapped stimuli, e.g. spatial distribution of the target [38]. We therefore measured performance, looking for interference, at asymptotic (short SOA) levels, with original and with swapped stimuli.

Fig. 4 demonstrates that following training with the original stimuli and then with the swapped stimuli, re-test with the original stimuli shows that performance threshold is (almost fully) retained and not hampered (threshold for re-test is 40 ± 2.4 ms, compared to original asymptote at 36 ± 2.3 ms; difference is not significant). We conclude that though the training effect is specific to the original orientations, as predicted by the earliest modification hypothesis, there is no interference between training with the original orientations to subject performance with swapped orientations—on the contrary, there is even some initial transfer for long SOAs. Furthermore, second training with the swapped orientations has no detrimental effect on re-test with the original orientations. We must therefore reject the earliest modification hypothesis, and accept the alternative output level modification hypothesis.

4.2. Floor effect

Could the lack of asymptotic interference between learning swapped and original orientations result from floor effects? That is, might there be differences between the system’s performance under the swapped and original stimuli conditions, but these are masked by

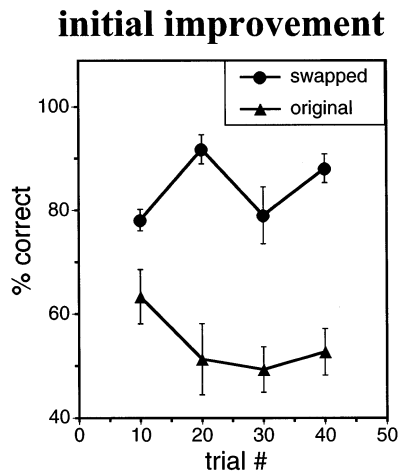


Fig. 5. Rate of success during the first 40 trials with original (triangles) and swapped (circles) target and distractor orientations. SOA for first 20 trials was 166 ms; for next 20 trials 116 ms. Error bars indicate inter-subject standard error. Note that with the original orientations performance is near chance, and there is even some early deterioration in performance, while for the swapped orientations, performance begins at a much improved level. Thus, there appears to be no initial confusion (for long SOA trials) in contrast to the prediction of the earliest modification hypothesis.

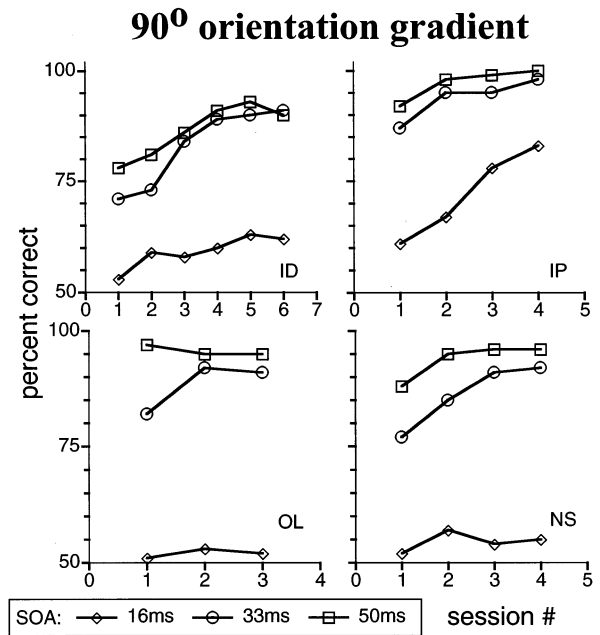


Fig. 6. Performance for easy task—90° difference between orientations of target and distractors. Performance is shown as percent correct for trials with different SOAs. There is nearly perfect performance for 33 ms SOA for all subjects so the threshold must be substantially below that. Thus, the similarity of thresholds for original and swapped orientations in Fig. 4 can not derive from a floor effect.

other limitations which keep the asymptotic threshold SOA as high as ~ 30 ms. Perhaps, had there not been additional limitations, we might have seen a lower threshold for one set of stimuli compared with the other set.

If such floor effects existed, it would mean that subjects could not reach an SOA lower than 30 ms with this behavioral paradigm. To ensure that this is not the case, we trained a group of four additional subjects with the same paradigm, only now target deviated by 90° from distractor elements. We figured that subjects should reach better asymptotic performance under these conditions if there is no floor effect. Indeed we found that, for all four subjects, threshold SOAs following learning (1–3 sessions were sufficient) were < 20 ms. As shown in Fig. 6, all performed nearly perfectly with 33 ms SOA, and two performed significantly above chance with 16 ms SOA. We conclude that the lack of interference at asymptote does not stem from a floor effect. It reflects a genuine co-existence of training-induced performance improvement for apparently conflicting sets of stimuli.

Perhaps the lack of interference derives from target detection being facilitated by enhanced responses of neighboring distractor detectors? If orientation salience is summed over a somewhat extended local area, the increased salience of neighboring elements could also explain the moderately improved performance follow-

ing swap compared to that of naive subjects. This explanation has two predictions, both of which we tested and refuted.

First, if salience were summed, then increasing the orientation gradient of the surrounding area would facilitate target detection. Thus, detection of a target surrounded by a ring of distractors with an intermediate orientation would be easier than detecting a target embedded in a homogeneous field of distractors. To test this, a group of 18 trained subjects was tested with successive sessions or interleaved blocks of these two conditions (target and distractor orientations: 30 and 60° versus target and distractors 30 and 60° with a ring of distractors surrounding the target with orientation of 45°). Consistently, performance was somewhat better under the single local gradient condition, compared with the surrounding ring condition (single target detection minus target-with-ring detection difference was $+7 \pm 4\%$ for successive sessions ($n = 6$) and $+5 \pm 1\%$ for interleaved blocks ($n = 12$); the difference was positive for all but one subject, for whom it was negligible). These results are consistent with previous findings indicating that salience is determined by very local orientation gradients [27].

The second prediction is that asymptotic performance with swapped orientations (and following re-swap) would be better than asymptotic performance on original orientations (before training with swapped orientations; compare Fig. 2 bottom, and middle, with top). However, our results show that in fact asymptotic performance following swapping is, if anything, somewhat worse than original asymptotic performance. Taken together, these two lines of results refute the alternative explanation.

In summary, it is not possible that performance improvement is largely due to asymmetric interactions between detectors for the target and the distractor orientations. Other mechanisms must be involved (at least under one stimulus condition) and these are sufficient to achieve nearly the same asymptotic level for both conditions. The earliest modification hypothesis must be rejected.

4.3. Spatial distribution of detection

An alternative interpretation to the lack of interference in asymptotic threshold, is that the threshold is an average across time (SOAs) and/or space (target at various array positions). While the average threshold has not been affected, its temporal and spatial components could have been. We thus asked, what is the pattern of spatial distribution for detection during initial, and asymptotic stages with original and with swapped orientations? If learning stems from the same lateral interaction mechanisms during first and second learning, asymptotic distribution of detection should

differ for original and swapped orientations. That is, where improvement was achieved for one set of orientations it should hamper performance with the other set, and in visual field locations where little improvement occurred, second learning should be expedited.

To test this prediction, we plot the initial and final spatial distributions with each of the two orientation sets. The initial and final 2-dimensional distributions with original orientations are plotted in Fig. 7 top (left and right, respectively). The initial distribution shows a large anisotropy: decay in detection is much steeper

spatial distribution of detection

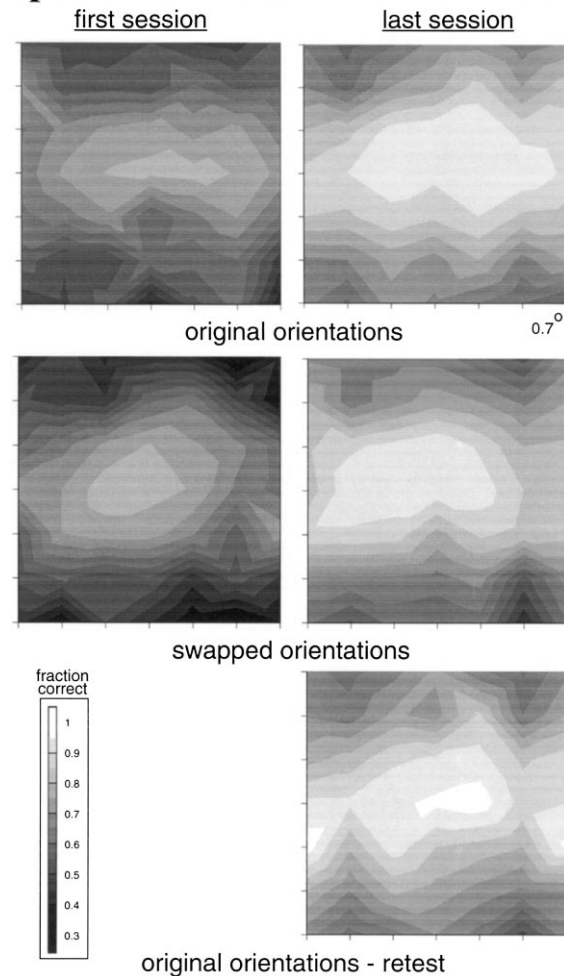


Fig. 7. Two dimensional distributions of detection with original (top and bottom) and swapped (middle) orientations. Left and right plots are averages for ten subjects for the first and last sessions, respectively. Note the expansion of the central bright (good performance) region with training, first for the original orientations, and then with the swapped orientations. Though there is some effect on the initial distribution with the swapped orientations due to the original training (in the area of central vision-middle left), the overall performance is similar to that of the naive subjects (top left). The final asymptotic distributions for the original and swapped orientations (top and middle right) are very similar, as are the distributions for original orientations before and after second learning with swapped orientations (top and bottom right). See text.

vertically compared with horizontally [39]. Learning raised detection for all positions (brightening of entire graph), with some expansion of the central region, mainly along the vertical meridian so that asymptotic performance is less anisotropic.

Fig. 7 middle row plots the 2-dimensional spatial distribution tested with swapped orientations. The middle left distribution is for performance on the first session with the swapped orientations following extensive training with the original stimuli, and the middle right distribution is for asymptotic performance following second learning with the swapped orientations. The performance distribution immediately following swap has a central cone that is less anisotropic—it is expanded vertically and is squeezed along the horizontal axis and is somewhat rotated towards the top-right to bottom-left diagonal. On the other hand, the decay towards the edges is steeper than the original one, as expressed by the darker ‘frame’ in the first session with swapped orientations. Following second training, the spatial distribution of detection with swapped orientations has a similar shape to that of asymptotic detection with original orientations. Thus, in this swapped-orientations case, improvement from first to last session was mainly at the stimulus’ outskirts.

Finally, Fig. 7, bottom, shows the detection distribution for the original orientations on a retest following the second swapped-orientations learning. Note that this distribution appears very similar to that of asymptotic performance before second learning (Fig. 7 top right). That is, second learning with swapped orientations did not interfere with the original learning effects. Furthermore, these distributions with the original orientations are similar to that for the swapped orientations, demonstrating the co-existence of improved mechanisms for seemingly conflicting stimulus situations.

In summary, the spatial distribution of the first test with swapped orientations is different than the naive distribution—there is some positive transfer of training effect to the swapped condition, rather than interference. An interesting spatial aspect is the spatial distribution of the transfer characteristics: improvement near fixation was transferred more than improvement at distal positions. This may result from easier learning near fixation being less specific than difficult learning at the periphery⁴ [21,38]. Yet, the total spatial improvement with respect to original first session is very similar for original and swapped orientations (i.e. comparing Fig. 7 top-left with top and middle right), as well as for original orientations before and after second learning (i.e. comparing Fig. 7 top-left with top and bottom right). Both these findings are in contrast to the predic-

time course of second learning

typical case (n=8):
2nd slower than original

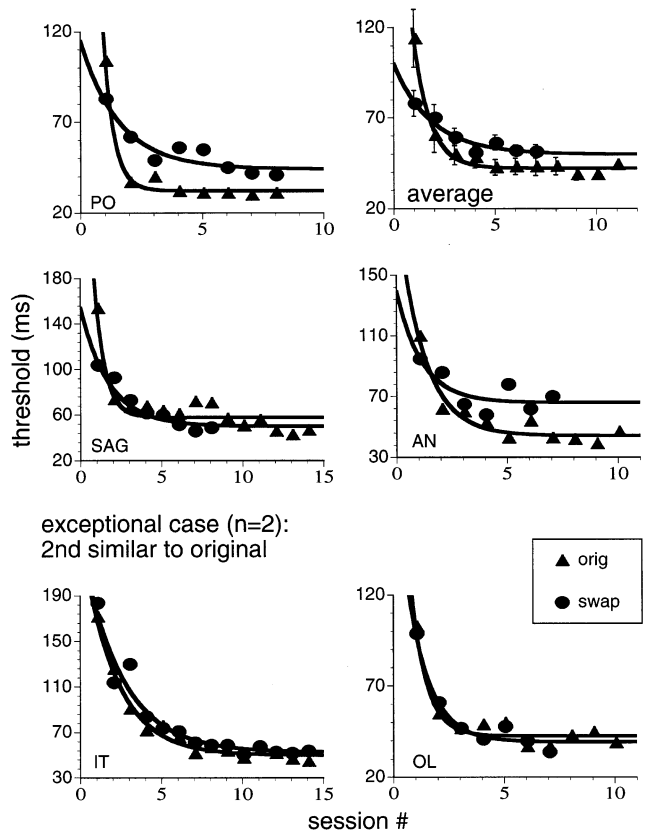


Fig. 8. Inter-subject differences. Two modes of second learning were seen: top four graphs are three examples and average for eight subjects (top right) whose second learning, with swapped orientations, started with small transfer and proceeded with impeded learning rate, as shown by the superimposed data points and curves. Some small difference in final asymptote is also apparent, but continued training might have led to identical performances. Bottom two graphs are learning curves of the two subjects whose first and second learning curves were very similar, both in terms of initial point and learning rate. Final data points shown with original orientations were measured following training with swapped orientations. The curves were computed by best fit to the exponential dependence: $\text{threshold} = a + b \cdot \exp(-c \cdot \text{session \#})$, where a (asymptotic threshold), b (difference between initial and final thresholds), and c (learning rate constant) are free parameters.

tions of the assumption of earliest modification site. Once again, the data lead us to the conclusion that the earliest modification hypothesis must be rejected in favor of the output level modification hypothesis.

4.4. Rates of learning and individual differences

Is the process of learning with swapped orientations independent of previous training or is it facilitated or impeded by prior experience? A crude comparison between the rates of the initial learning and learning with

⁴ Note that the stimulus elements were not scaled with eccentricity.

swapped orientations also may be seen in Fig. 4. Interestingly, although the threshold of the first session with modified stimuli was lower than the initial threshold with the original stimuli, the threshold for the second session with swapped orientations was higher than the threshold for the second session with the original orientations. This already indicates that despite there being some transfer in performance, there may be a concomitant reduced rate of subsequent learning.

Although the across-subject average shows moderate transfer together with impeded subsequent learning rate, this pattern was not found for all subjects. These individual differences are demonstrated in Fig. 8. Eight of the ten subjects showed a slower learning curve with swapped orientations. Fig. 8 top shows three examples of this pattern of slowed learning, as well as the average for these eight subjects. First and second learning are superimposed to allow direct comparison. In addition to impeded learning rate, these subjects (except one—EI in Fig. 4) also showed a small degree of transfer. Thus, for most subjects, though there was no interference for asymptotic performance, there was some slowing of the learning itself: though different mechanisms must underlie performance with different orientation sets, there may be some interaction between them during the initial second learning stages.

For two subjects (out of ten), however, second learning, following orientation swapping, seemed like a ‘replay’ of first learning. That is, performance was not transferred and learning rate was not hampered. First and second learning curves of these two subjects are plotted in Fig. 8, bottom. We do not have a sufficiently large sample of subjects to determine if these are distinct groups, or simply ends of a continuum of learning patterns. Similarly, we found this slowing of second learning disappears for all subjects under other conditions [21].

In summary, the pattern of transfer and interference we reported here shows that swapping orientations does not cause the expected initial confusion. There is orientation specificity, and even some transfer for easy cases. We found co-existing, similar, non-interfering, asymptotic performance abilities for original and swapped orientations. These characteristics are general across the tested paradigms and training procedures. We conclude that the ‘earliest modification’ hypothesis must be rejected in favor of the alternative ‘output level modification’ hypothesis.

5. Discussion

5.1. Learning mechanisms

The basic finding of this study is that there was no interference between asymptotic performance with orig-

inal and with swapped target and distractor orientations (as demonstrated in Figs. 4, 7 and 8). In fact, there is a small, but definite degree of positive transfer to first (easy) trials with swapped orientations. This finding strongly rejects the whole family of models based on the assumption of changes within lateral interactions between detectors for target and distractor orientations, as predicted by the hypothesis of earliest modification site (with the assumptions specified in Footnote 2). The only way the cortex can maintain contradictory relationship performance asymptotes is to base the training improvement on computations performed with the early site’s output. Thus, we conclude that any ‘strong version’ of the earliest modification hypothesis (asserting that most of the behavioral improvement stems from modifications at an early site and the same strategies are applied for first and second learning) must be rejected. Rather, we propose that the training produces improved computations in subsequent, higher cortical levels, where further computations are performed using the earlier site’s output. Here, complementary, and even contradictory computations may co-exist, without interfering with each other, as would be the result of modifying lower site reciprocal lateral interactions. Thus, while separate representations for the target-present versus target-absent categories are probably formed already at the V1 level (see Section 1), learning induced modifications mostly occur at V1 output levels. The latter conclusion stems from a detailed study of stimulus specificities when learning pop-out detection [10]. While specificities are compatible with an early site, the transfer found for target shift across mid-line ($> 1^\circ$) can probably not stem from V1 receptive field sizes. The actual area at which modifications occur may largely depend on training conditions [21].

We have seen indications here of another robust effect that is not easily reconciled with the earliest modification hypothesis, an effect that we demonstrated in detail in a separate paper [21]. Improvement transfers across orientation manipulations for easy trials (when the target appears near fixation, when SOA is long, and/or when target-distractor orientation difference is large) but not for difficult trials. The earliest modification assumption does not predict a dependence of the relative transfer on the degree of trial difficulty. On the contrary, modification of the earliest orientation detectors should affect detection under all conditions, easy or hard, since all should depend on these same detectors. On the other hand, a higher level of parallel computation could accommodate such dependence.

Additional support for the output level hypothesis⁵ is the task dependence of perceptual learning found for this task [29] and for other tasks [9,40].

⁵ We thank an anonymous reviewer for this comment.

We conclude that learning results mainly from improved selection of task related inputs to secondary cortical levels. We have called this the ‘output level modification’ hypothesis, which accordingly we adopt.

5.2. Relation to previous behavioral findings

Schneider and Shiffrin [41] and Shiffrin and Schneider [42], studied the performance of subjects trained on visual search for a target embedded in distractors, comparing performance with fixed or reversing target/distractor conditions. They trained subjects either with a consistent target (e.g. digits) embedded in a field of distractors (e.g. letters), or with variable (within session) mapping conditions, so that the target of one trial could be the distractor of another, and vice versa. They found that under consistent stimulus-response mapping, performance is automatized (becomes independent of number of distractors), while in the variable case, there is less improvement and no automatization. Indeed, they did not find substantial stimulus specificity for the variable case [43]. Shiffrin and Schneider [42], also tried consistent condition training followed by training with swapped conditions. They found that performance immediately following the switch was worse than naive performance. The findings of our study show that subjects can be trained for optimal performance with swapped stimuli without negative transfer. The reason sequential training is efficient in our case may be that pop-out is initially automatic, and its automaticity does not depend on training. In addition, it appears that subjects consistently search not for a single target element (with a particular orientation), but rather for a more complex structure (e.g. target surrounded by distractors, see Fig. 3). With regard to this complex target structure, the mechanisms used for the two sets of stimuli (original and swapped) may not be inconsistent, but orthogonal.

5.3. Comparison with previous concepts

In this study we have shown that learning does not mainly result from modification of any type of target-distractor lateral interactions. In a more general context, we infer that learning occurs mainly at an output selection level, rather than by setting up new early-level perceptual modules with additional examples [8,19]. However, the alternative extreme, that the system actively seeks basic organizational principles, as suggested by E. Gibson’s theory of perceptual learning, [44], is also not supported by the data. According to her theory, pop-out improvement would result from learning which characteristics are most reliable for detecting large orientation differences within the array of stimulus elements. This learning should transfer to other orientations, positions, etc. If conditions do not allow

for complete transfer, the degree of transfer should at least increase with practice. Thus, the prediction of such a hypothesis is that learning should be less specific as it proceeds towards attaining its internal goal. We found no evidence supporting this concept. If anything, it seems that learning becomes more specific as a function of practice (Ahissar and Hochstein [11], and the present paper; Karni and Sagi [45], and for a review Ahissar and Hochstein [32]).

A less demanding hypothesis is that learning is generalized beyond the actual trained examples, but not beyond sub-categories. In the pop-out case, a category could be a specific target-distractor relationship. For example, learning to detect a single element which has an orientation which is ‘more vertical’ than that of its neighbors. This would explain the large deficit following orientation swapping (above, Section 4, Fig. 4) or rotation by 90°, on the one hand, and transfer across mirror image reversal, on the other [11]. However, we previously found that rotation by only 30° (15–45° to 45–75°, or vice versa [11]) greatly degraded performance. Thus, this model, too, is not consistent with characteristics of learning pop-out detection.

A selection model that learns specific examples was suggested by Seung and Sompolinsky [46]. In their model, improvement in orientation discrimination originated from an improved exploitation of information within orientation selective units by increasing the weight assigned to the most informative units. Optimization was sought only at the highest, task-specific stage. Weiss et al. [47] also devised a model where improvement (in a vernier hyperacuity task) stems from modifications in output connections⁶. Treisman and her colleagues reached a similar conclusion [40,48]. They found that training to detect a complex target yields improved performance, but does not produce a new complex feature. That is, performance of other tasks using the same target stimulus is not improved. These concepts are in agreement with our findings.

5.4. Relation to physiological findings

Physiological results stress the expansion of cortical representation of the specific range of parameters that were consistently enhanced by the experimental paradigm. However, the direct relevance of these modifications to perceptual learning has not been shown [49]. Moreover, these modifications will not yield improvement unless the readout mechanism will be appropriately updated. On the other hand, improvement may be achieved by optimizing the readout mechanism with no change of basic feature representations. The physio-

⁶ However, in this model they used an Exposure-Dependent-Learning rule that is incompatible with the previous behavioral findings of Ahissar and Hochstein [10].

logical correlate of modifications within a readout mechanism is not known and thus hard to test. Ahissar and Ahissar [49], suggest that it may be found within response properties of the same neurons, at longer time intervals with respect to stimulus onset. However, no comparison between modifications within transient initial response properties and subsequent (sustained) responses was systematically applied. It thus seems that current physiological findings cannot eliminate either possibility. Naturally, not all improvement should derive from a similar type of mechanism. However this question was not resolved for any learning paradigm.

We conclude that the ‘earliest modification’ option could not account for behavioral improvement resulting from training pop-out detection. Rather, improvement stems mainly from modifications at subsequent levels receiving the output of the initial cue-discriminative representation.

Acknowledgements

We thank Robert Shapley for important suggestions in the course of this study, for reviewing manuscript drafts, and suggesting the control experiment of Fig. 6. We thank Anne Treisman for collaboration, counsel and comments on this work. We thank an anonymous reviewer for comments leading to the ‘ring’ control experiments. The research was supported by grants from the US-Israel Binational Science Foundation (BSF) and the Israel Science Foundation of the Israel Academy of Arts and Sciences.

References

- [1] Ramachandran VS, Braddick O. Orientation specific learning in stereopsis. *Perception* 1973;2:371–6.
- [2] Fiorentini A, Berardi N. Perceptual learning is specific for orientation and spatial frequency. *Nature* 1980;287:43–4.
- [3] Ball K, Sekuler R. A specific and enduring improvement in visual motion discrimination. *Science* 1982;218:697–8.
- [4] Ball K, Sekuler R. Direction specific improvement in motion perception. *Vis Res* 1987;27:953–65.
- [5] Nazir T, O’Regan JK. Some results on translation invariance in the human visual system. *Spat Vis* 1990;5:81–100.
- [6] Karni A, Sagi D. Where practice makes perfect: evidence for primary visual cortex plasticity. *Proc Natl Acad Sci USA* 1991;88:4966–70.
- [7] O’Toole A, Kersten D. Learning to see random dot stereograms. *Perception* 1992;21:227–43.
- [8] Poggio T, Fahle M, Edelman S. Fast perceptual learning in visual hyperacuity. *Science* 1992;256:1018–21.
- [9] Shiu L, Pashler H. Improvement in line orientation discrimination is retinally local but dependent on cognitive set. *Percept Psychophys* 1992;52:582–8.
- [10] Ahissar M, Hochstein S. The role of attention in early perceptual learning. *Proc Natl Acad Sci USA* 1993;90:5718–22.
- [11] Ahissar M, Hochstein S. Learning pop-out detection: specificities to stimulus characteristics. *Vis Res* 1996;21:3487–500.
- [12] Polat U, Sagi D. Spatial interactions in human vision: from near to far via experience dependent cascades of connections. *Proc Natl Acad Sci USA* 1994;91:1206–9.
- [13] Beard BL, Levi DM, Reich LN. Perceptual learning in parafoveal vision. *Vis Res* 1995;35:1679–90.
- [14] Schoups AA, Vogels R, Orban GA. Human perceptual learning in identifying the oblique orientation: retinotopy, orientation specificity and monocularly. *J Physiol* 1995;483(3):797–810.
- [15] Fahle M, Edelman S, Poggio T. Fast perceptual learning in hyperacuity. *Vis Res* 1995;35:3003–13.
- [16] Rubin N, Shapley R, Nakayama K, Grossetete A. Abrupt learning in illusory contour perception. *Investig Ophthalmol Vis Sci* 1996;35(4):771.
- [17] Rubin N, Nakayama K, Shapley R. Abrupt learning and retinal size specificity in illusory-contour perception. *Curr Biol* 1997;7:461–7.
- [18] Desimone R, Ungerleider LG. Neural mechanisms of visual processing in monkey. In: Boller F, Gafman J, editors. *Handbook of Neuropsychology*, vol. 2. Amsterdam: Elsevier Science Publishers, 1989:267–299.
- [19] Poggio T, Edelman S, Fahle M. Learning in visual modules from examples: a framework for understanding adaptive visual performance. *CVGIP: Image Understanding* 1992;56:22–30.
- [20] Peres R, Hochstein S. Modeling perceptual learning with multiple interacting elements: a neural network model describing early visual perceptual learning. *J Comput Neurosci* 1994;1:323–38.
- [21] Ahissar M, Hochstein S. Task difficulty and the specificity of perceptual learning. *Nature* 1997;387:401–6.
- [22] Hubel DH, Wiesel TN. Receptive fields and functional architecture of monkey striate cortex. *J Physiol* 1968;148:574–91.
- [23] Treisman A, Gelade G. A feature integration theory of attention. *Cogn Psychol* 1980;12:97–136.
- [24] Julesz B. Textons, the elements of texture perceptions and their interactions. *Nature* 1981;290:91–7.
- [25] Merigan WH, Nealy TA, Maunsell JHR. Visual effects of lesions of cortical area V2 in macaques. *J Neurosci* 1993;13:3180–91.
- [26] Knierim JJ, Van Essen DC. Neuronal responses to static texture patterns in area V1 of the alert Macaque monkey. *J Neurophysiol* 1992;67:961–980.23.
- [27] Nothdurft HC. Sensitivity for texture gradient in texture discrimination tasks. *Vis Res* 1985;25:1957–68.
- [28] Sagi D, Julesz B. Short range limitation on detection of feature differences. *Spat Vis* 1987;2:39–49.
- [29] Ahissar M, Hochstein S. Learning is both stimulus and task specific. *Investig Ophthalmol Vis Sci* 1991;32:250.
- [30] Ahissar M, Hochstein S. Spatial anisotropy in feature search performance and learning. *Investig Ophthalmol Vis Sci* 1994;35(4):3818.
- [31] Ahissar M, Hochstein S. How early is early vision? Evidence from perceptual learning. In: Papanthomas TV, editor. *Early Vision and Beyond*. Cambridge: M.I.T. Press, 1995:199–206.
- [32] Ahissar M, Hochstein S. Perceptual learning. In: Walsh V, Kulikowski J, editors. *Perceptual Constancy: Why Things Look as They Do*. Cambridge: Cambridge University Press, 1998:455–98.
- [33] Koch C, Ullman S. Shifts in selective visual attention: towards the underlying neuronal circuitry. *Hum Neurobiol* 1985;4:219–27.
- [34] Blakemore C, Tobin A. Lateral inhibition between orientation detectors in the cat’s visual cortex. *Exp Brain Res* 1972;15:439–40.
- [35] Grinvald A, Lieke EE, Frostig RD, Hildesheim R. Cortical point-spread function and long-range lateral interactions revealed by real-time optical imaging of macaque monkey primary visual cortex. *J Neurosci* 1994;14:2545–68.

- [36] Wolfe JM, Friedman-Hill SR, Stewart MI, O'Connell KM. The role of categorization in visual search for orientation. *J Exp Psychol Hum Percept Perform* 1992;18:34–49.
- [37] Quick RF. A vector magnitude model of contrast detection. *Kybernetik* 1974;16:65–7.
- [38] Ahissar M, Hochstein S. Perceptual learning transfer over space and orientation. *Investig Ophthalmol Vis Sci* 1996;37(3):3182.
- [39] Krose BJA, Julesz B. The control and speed of shifts of attention. *Vis Res* 1989;29:1607–19.
- [40] Treisman A, Vieira A, Hayes A. Automaticity and pre-attentive processing. *Am J Psychol* 1992;105:341–62.
- [41] Schneider W, Shiffrin RM. Controlled and automatic human information processing. I. Detection, search, and attention. *Psychol Rev* 1977;84:1–66.
- [42] Shiffrin RM, Schneider W. Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychol Rev* 1977;84:127–90.
- [43] Sireteanu R, Rettenbach R. Perceptual learning in visual search: fast, enduring, but non specific. *Vis Res* 1995;35:2037–43.
- [44] Gibson EJ. Principles of perceptual learning and development. New York: Appleton-Century-Crofts, 1969.
- [45] Karni A, Sagi D. The time course of learning a visual skill. *Nature* 1993;365:250–2.
- [46] Seung HS, Sompolinsky H. Simple models for reading neuronal population codes. *Proc Natl Acad Sci USA* 1993;90:10749–53.
- [47] Weiss Y, Edelman S, Fahle M. Models of perceptual learning in vernier hyperacuity. *Neural Comput* 1993;5:695–718.
- [48] Treisman A. Perceiving and re-perceiving objects. *Am Psychol* 1992;47:862–75.
- [49] Ahissar E, Ahissar M. Plasticity in auditory cortical circuits. *Curr Opin Neurobiol* 1994;4:580–7.