Perceptual learning: learning to see

Dov Sagi and David Tanne

The Weizmann Institute of Science, Rehovot, Israel

Perceptual learning in vision has been found to be highly specific for simple stimulus attributes, implying highly specific modifications in the nervous system. The type of specificity found (location, orientation, eye) implied plasticity at very early stages of visual processing, where processing modules were believed to be hard-wired and task independent. Recent studies show, however, that learning is task dependent. Studies examining the time course of learning indicate that at least two different learning processes are involved in perceptual learning, reflecting different levels of processing. Perceptual learning appears to be governed by associative rules and to be constrained by system architecture.

Introduction

When a person is asked to perform a visual (or any other sensory) discrimination task, it is often the case that they improve with practice, even on very simple tasks. This improvement occurs without any reinforcement and does not seem to involve conscious effort, but rather it seems to be controlled by some inherent subconscious process. During the past decade, perceptual learning has been shown to be involved in a variety of visual tasks, such as stereoscopic vision [1,2], gratings detection [3,4], hyperacuity [5,6], phase discrimination [7], motion detection [8], texture discrimination [9,10], search [11], and pattern discrimination [12]. Some of these studies showed specificity of learning for location in the visual field (i.e. what is learned at one location cannot be used when stimulus is presented at another location) [2,6–8,10,12], for orientation [1,4,6–8,10], for spatial frequency [7] and for direction of motion [8]. In some cases, learning was found to persist for a few weeks without further practice [7,8].

In this review we will discuss the results from a new wave of studies, which are providing more insight into the processes involved in perceptual learning. In particular, studies on texture segmentation, lateral masking and hyperacuity are generating interesting results. Like the earlier studies, these experiments demonstrate the specificity of learning for stimulus features. However, learning was also found to be specific for non-stimulus variables, such as the task used for training, implying higher level controls over learning. These studies also demonstrated the existence of two types of learning, fast (binocular) and slow (monocular). The slow phase requires a consolidation period of about six hours, which, when during sleep time, depends on the

integrity of the REM (rapid eye movement) sleep stage. We will also discuss results from recent lateral masking experiments, which have been shown to support an associative learning model for perceptual learning.

Stimulus-driven versus task-driven learning

One important development in understanding perceptual learning is the finding that learning involves modification of the sensory representation in the brain [10]. Thus, perceptual learning is not only a way of training attention to pick up distinctive stimulus features [13] or of improving sensory processing resulting from increased alertness [14]. Karni and Sagi [10] studied learning of orientation-based texture segmentation. This task is believed to be carried out by a pre-attentive (parallel across the visual field, bottom-up) stage of visual processing that detects texture boundaries by comparing activities of neighboring local processing units (filters) selective for basic image features, such as orientation and size [15]. Single-cell recordings show that texture boundaries are already detectable in cortical area V1 [16]. In the learning experiment, observers had to identify the orientation of a small texture region embedded in a background of a different orientation [10]. An additional fixation task was used concurrently with the texture task to minimize the use of attention [17] in the segmentation process. Learning was found to be spatially local (i.e. no transfer of learning between two retinal locations separated by three degrees), specific to background orientation and 'monocular', thus supporting the idea that learning takes place at a low-level anatomical site. Monocular learning is difficult to reconcile with the 'attention' hypothesis [13], as attentive

---

Abbreviation
REM—rapid eye movement.

processes cannot make a selection based on eye-of-origin information [18].

Ahissar and Hochstein [19••] also used a texture stimulus (the target, however, was only one line segment), but trained the observers on two different tasks. In one task, observers had to detect the target (i.e. a line that was oriented differently from the others), and in the second task, they had to identify the shape outlined by the outer borders of the background. The authors [19••] found almost no transfer of learning between the tasks, though the stimuli used were identical in both tasks. Shiu and Pashler [20] also found that orientation discrimination does not improve when observers practice on brightness discrimination using the same lines as stimuli. It was found, however, that there exists partial transfer between color discrimination tasks and vernier acuity tasks, and also between eyes when using these tasks (P Moller, T Crabb, AC Hulbert, *Perception* 1993, 22:37). These results imply that perceptual learning affects a selected population of sensory processing units, and that these units are probably selected by attention.

Note that the different tasks involve different low-level processing units (e.g. brightness discrimination may be better performed by non-oriented units when target orientation is not known), and it is possible that the system selects (or attends to) the most efficient units for the task in order to optimize performance. As this optimization amounts to detailed feedback input into the sensory processing units, it is possible that what is learned is also the use of feedback. Learning may take place at all processing stages and, therefore, it may be possible to study a particular stage by selecting the right task.

Whereas pre-attentive texture-segmentation tasks probably probe primary visual areas, attentive search task may be useful for exploring higher-level stages of processing. A recent report suggests that learning an attentive search task is spatially local, but can be both retinotopic- and object-centered (S Suzuki, P Cavanagh, *Invest Ophthal Vis Sci Abstr* 1993, 34:1233). It was also reported that attentive serial search can become parallel after practice, but this practice effect is not stimulus specific (R Sireteanu, R Rettenbach, *Perception* 1993, 22:36). So it seems that attentive search tasks involve some space-invariant, and stimulus-independent, learning.

Another related issue is the role of response feedback. It was found that feedback is not necessary for learning [8]. Shiu and Pashler [20] compared learning curves under three conditions: trial-by-trial feedback, block-by-block feedback and no feedback. They did not find any significant difference between the first two conditions. On the third condition, no feedback, observers did not show much improvement within each session, but rather improved from one session to the next. Using a hyperacuity task, Fahle and Edelman [21•] found that learning without feedback is slightly

slower than with feedback. These findings rule out any model of learning that is based on detailed supervision — in which the internal response pattern is compared with an expected one in order to modify connectivity. It is still possible that observers have some indication of the correctness of their response when stimuli are clearly discriminable and, thus, an internal feedback can be used to speed up learning [22]. This assumption is supported also by the informal observation that though observers can perform well after practice on initialy impossible discrimination tasks, they can not reach this performance level unless they are trained first on easy conditions.

Time course of learning

Improvement of performance may take place between trials, between blocks or between sessions, with time scales ranging between seconds to hours. It appears that at least two processes are involved [23••]. Observers practicing on a texture segmentation task [10] showed an initial fast saturating phase of learning, followed by a slow phase where improvement could be seen between sessions. The fast phase showed interocular transfer of learning, whereas the slow phase was found to be monocular, thus implying that two different stages of processing are involved in the learning process. Both stages failed to show learning transfer across locations and orientations. This new dichotomy can shed light on the earlier results of Fiorentini and Berardi [7] on phase discrimination where learning showed fast saturation (200 trials, within a single session) and complete interocular transfer, and of Ball and Sekuler [8] on motion discrimination where learning was somewhat slower (3–4 sessions for saturations) and interocular transfer was incomplete (74 %).

Retention of texture learning was found to be complete after 2–3 years [23••], a new record as compared with ten weeks for motion discrimination [8], and between six weeks to six months (partial retention) for phase discrimination [7].

Karni and Sagi [23••] also found that between-session improvement occurs only if the two sessions are separated by at least 6–8 hours, suggesting a consolidation period. Consolidation was found to take place during the awake state (normal life activity) and sleep time. When consolidation occurred during sleep, it was found to be dependent on the integrity of the REM sleep stage (A Karni, D Tanne, BS Rubenstein, JJM Askenasy, D Sagi, *Soc Neurosci Abstr* 1992, 18:387). Subjects did not show learning effects when the consolidation period was constrained to sleep time and their REM stage sleep was disrupted, though their normal performance level on trained tasks was retained. Slow-wave stage sleep deprivation had only a minor effect on learning.

Learning rules

Learning may involve improvement of stimulus–response connections and/or detection of associations (correlated activities) within the sensory system. Some insight into the mechanism of learning comes from lateral masking experiments. Polat and Sagi [24] found that contrast threshold for an oriented Gabor signal can be enhanced by positioning two high-contrast flanking signals (masks) at a distance of about three times the target wavelength. This enhancement was found only when the masks were placed in the direction defined by the target orientation or orthogonal to it (cardinal directions) [25]. Learning experiments showed that practice can increase the range of these interactions by a factor of six, but only along the cardinal directions [26••]. An increase in the range could not be obtained by practicing on the large distances alone, rather, it required practicing with a mixture of distances, including the small ones. The finding that different distances have to be processed within the same session implies that there exists a window of integration time for the activities of the different neurons to be detected and associated. 'Un-learning' was also observed when practice was limited to a small range of distances. Polat and Sagi [26••] suggested that long-range interactions are generated via cascades of local interactions, with local connections being strengthened or weakened by practice. Learning was suggested to be limited to connections that already exist in the system (cardinal directions) and to follow associative (Hebb-like) rules. According to this theory, activation of two neighboring units within a short time interval (of a few minutes) increases the efficacy of their connections (if they exist), whereas repetitive activation of one of them reduces connection efficacy.

Electrophysiological correlates

Though the functional organization of primary sensory areas was considered to be fixed in the adult primate, extensive work done in the past decade has revealed that a great deal of plasticity exists in those areas (for reviews see [27–30]). Long-term changes in neural responses can be induced by behavioral training in different sensory modalities. Recanzone, Merzenich and colleagues [31,32] demonstrated major changes in the details of the cortical spatio-temporal representation in monkeys (adult owl) trained to detect tactile stimuli delivered to a small fixed spot on one finger. Progressive improvement in auditory frequency discrimination with practice appeared to be accounted for by a parallel, progressive change in the region representing the corresponding frequencies in the primary auditory cortex [33••]. Weinberger and associates [34,35•] demonstrated frequency-specific receptive field plasticity in auditory cortical areas by studying classical conditioning in adult guinea pigs. A recent report suggests Hebb-like rules apply to this type of learning (SJ

Cruikshank, NM Weinberger, *Soc Neurosci Abstr* 1993, 19:164).

Several recent studies show plasticity without behavior. Fregnac, Shulz and colleagues [36,37] describe a cellular analog of primary visual cortex plasticity. They induced long-term changes in receptive field selectivity (for orientation and eye), in anesthetized and paralyzed cats, by pairing specific stimuli with selective activation (iontophoretically) of the recorded neuron. Although these stimuli were not behaviorally significant, driving the cells iontophoretically probably provided a simulated gating signal. In the primary visual cortex, a striking increase in receptive field size was observed after removing visual input by using restricted retinal lesions (long term) [38,39] or after producing an 'artificial scotoma' by masking out an area covering the receptive field of a recorded neuron (short term) [40]. These visual effects were found in anesthetized and paralyzed cats. Single-units recordings from areas MT (medial temporal) and MST (medial superior temporal) of behaving macaque monkeys showed fast (300–500 trials) sensitivity improvement, and the rate of improvement correlates with the psychophysical performance (E Zohary, S Celebrini, KH Britten, WT Newsome, *Soc Neurosci Abstr* 1993, 19:1282). This improvement in neuronal sensitivity could also be obtained by repetitive stimulus presentation. In a recent study, electrical brain activity (VEP: visual evoked potentials) was recorded from the occipital areas while the observers were exposed to vernier stimuli, i.e. no task (M Fahle, *Perception* 1993, 22:30). Some specific field potential configurations appeared after a few hundreds of trials, presumably correlated with learning observed using the same stimuli [21•].

Conclusions

Psychophysical and electrophysiological studies provide evidence for plasticity of primary sensory areas. Perceptual learning seems to provide an excellent behavioral paradigm for exploring human learning, especially when coupled with current understanding of human vision. Perceptual learning was shown to have two major components: fast (few hundreds of trials) and slow (days). The fast component seems to affect higher levels of processing (above the site of binocular integration), and probably involves top-down processes, improving the link between task-dependent units and sensory units while selecting optimal sensory units for the task. Once these links become efficient, the task becomes 'automatic' (i.e. non-attentive [17]) and performance is then limited by sensory architecture only. The slow component seems to follow the fast one and involves low-level processes (monocular) within primary sensory areas. At this stage, links between sensory units are strengthened or weakened according to their activity correlations, thus establishing new associations and dissociations. First-order associations (via one link) are limited to direct connections,

but higher-order associations are possible by establishing chains of associations. Thus, although we are limited by the system architecture in what we can perceive as direct associations, it is possible to perceive more complex percepts (or concepts) by indirect associations. Fast learning probably takes place on-line, when the stimulus is still effective or immediately after, but slow learning and the consolidation of associations seem to be performed off-line, for hours after stimulus presentation while the individual is not aware of the problem being solved. Though consolidation of associations may take place during daytime, it seems that it also depends on processes that are active during REM (dream) sleep.

References and recommended reading

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

1.   RAMACHANDRAN VS, BRADDICK O: Orientation-Specific Learning in Stereopsis. *Perception* 1973, 2:371–376.

2.   O'TOOLE AJ, KERSTEN DJ: Learning to See Random-Dot Stereograms. *Perception* 1992, 21:227–243.

3.   DE VALOIS KK: Spatial Frequency Adaptation Can Enhance Contrast Sensitivity. *Vision Res* 1977, 17:1057–1065.

4.   MAYER MJ: Practice Improves Adults' Sensitivity to Diagonals. *Vision Res* 1983, 23:547–550.

5.   MCKEE SP, WESTHEIMER G: Improvement in Vernier Acuity with Practice. *Percept Psychophys* 1978, 24:258–262.

6.   POGGIO T, FAHLE M, EDELMAN S: Fast Perceptual Learning in Visual Hyperacuity. *Science* 1992, 256:1018–1021.

7.   FIORENTINI A, BERARDI N: Learning in Grating Waveform Discrimination: Specificity for Orientation and Spatial Frequency. *Vision Res* 1982, 21:1149–1158.

8.   BALL K, SEKULER R: Direction-Specific Improvement in Motion Discrimination. *Vision Res* 1987, 27:953–965.

9.   GURNSEY R, BROWSE RA: Micropattern Properties and Presentation Condition Influencing Visual Texture Discrimination. *Percept Psychophys* 1987, 41:239–252.

10.   KARNI A, SAGI D: Where Practice Makes Perfect in Texture Discrimination: Evidence for Primary Visual Cortex Plasticity. *Proc Natl Acad Sci USA* 1991, 88:4966–4970.

11.   STEINMAN SB: Serial and Parallel Search in Pattern Vision? *Perception* 1987, 16:389–398.

12.   NAZIR TA, O'REGAN JK: Some Results on Translation Invariance in the Human Visual System. *Spatial Vision* 1990, 5:81–100.

13.   GIBSON EJ: *Principles of Perceptual Learning and Development.* New York: Appleton-Century-Crofts; 1969.

14.   WOLFORD G, MARCHACK F, HUGHES H: Practice Effects in Backward Masking. *J Exp Psychol [Hum Percept Perf]* 1988, 14:101–112.

15.   RUBENSTEIN BS, SAGI D: Spatial Variability as a Limiting Factor in Texture Discrimination Tasks: Implications for Performance Asymmetries. *J Opt Soc Am [A]* 1990, 7:1632–1643.

16.   VAN ESSEN DC, DEYOE EA, OLAVARRIA JF, KNIERIM JJ, FOX JM, SAGI D, JULESZ B: Neural Responses to Static and Moving Texture Patterns in Visual Cortex of the Macaque Monkey. In *Neural Mechanisms of Visual Perception.* Edited by Lam DMK, Gilbert C. The Woodlands, Texas: Portfolio Publishing; 1989.

17.   BRAUN J, SAGI D: Vision Outside the Focus of Attention. *Percept Psychophys* 1990, 48:45–58.

18.   WOLFE JM, FRANZEL SL: Binocularity and Visual Search. *Percept Psychophys* 1988, 44:81–93.

19.   AHISSAR M, HOCHSTEIN S: Attentional Control of Early Perceptual Learning. *Proc Natl Acad Sci USA* 1993, 90:5718–5722.
••
Observers were briefly presented with textures composed of equi-oriented line segments in which a target line having a different orientation was embedded. One group of observers was trained to detect the target (local task) and another group was trained to identify the orientation of the rectangular shape outlined by the texture background (global task). Learning was found to be specific for the tasks tested and for task-related stimulus properties, such as local orientation for the local task and display size for the global task.

20.   SHIU L-P, PASHLER H: Improvement in Line Orientation Discrimination is Retinally Local But Dependent on Cognitive Set. *Percept Psychophys* 1992, 52:582–588.

21.   FAHLE M, EDELMAN S: Long-Term Learning in Vernier Acuity: Effects of Stimulus Orientation, Range and of Feedback. *Vision Res* 1993, 33:397–412.
•
Learning to detect a fine offset between lines was found to be orientation specific but not (offset) size specific. Absence of response feedback had only a slight effect on the learning rate.

22.   WEISS Y, EDELMAN S, FAHLE M: Models of Perceptual Learning in Vernier Hyperacuity. *Neural Computation* 1993, 5:695–718.

23.   KARNI A, SAGI D: The Time Course of Learning a Visual Skill. *Nature* 1993, 365:250–252.
••
Observers practicing a texture segmentation task [10] showed fast learning within the first session but constant performance within the subsequent sessions; however, performance improved between sessions if they were separated by at least six hours. Fast learning showed interocular transfer, whereas slow (between sessions) learning was found be monocular. Learning effects were retained for at least 2–3 years without further practice.

24.   POLAT U, SAGI D: Lateral Interactions Between Spatial Channels: Suppression and Facilitation Revealed by Lateral Masking Experiments. *Vision Res* 1993, 33:993–997.

25.   POLAT U, SAGI D: The Architecture of Perceptual Spatial Interactions. *Vision Res* 1994, 34:73–78.

26.   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, 91:1206–1209.
••
Observers were trained to detect a Gabor signal in the presence of two flanking high contrast but otherwise similar signals (masks). Target detection was found initially to only improve with the presence of masks at a short distance [24], implying short-range interactions. Practice had the effect of a continuous increase of interaction range, in a way that long-range interactions could not be established before medium-range interactions were becoming effective. It was suggested that long-range interactions are produced by chains of local interactions. Learning was found to be specific for eye, spatial frequency and retinal location (to within a neighborhood of half a degree).

27.   MERZENICH MM, SCHREINER C, JENKINS W, WANG X: Neural Mechanisms Underlying Temporal Integration, Segmentation and Input Sequence Representation: Some Implications for the Origin of Learning Disabilities. *Ann NY Acad Sci* 1993, 182:1–22.

28.   MERZENICH MM, SAMESHIMA K: Cortical Plasticity and Memory. *Curr Opin Neurobiol* 1993, 3:187–196.

29.   WEINBERGER NM: Learning-Induced Changes of Auditory Receptive Fields. *Curr Opin Neurobiol* 1993, 3:570–577.

30. GILBERT CD: Rapid Dynamic Changes in Adult Cerebral Cortex. *Curr Opin Neurobiol* 1993, 3:100–103.

31. RECANZONE GH, JENKINS WM, HRADEK GH, MERZENICH MM: Progressive Improvement in Discriminative Abilities in Adult Owl Monkeys Performing a Tactile Frequency Discrimination Task. *J Neurophysiol* 1992, 67:1015–1030.

32. RECANZONE GH, MERZENICH MM, SCHREINER CE: Changes in the Distributed Temporal Response Properties of SI Cortical Neurons Reflect Improvements in Performance on a Temporally Based Tactile Discrimination Task. *J Neurophysiol* 1992, 67:1071–1091.

33. RECANZONE GH, SCHREINER CE, MERZENICH MM: Plasticity in
•• the Frequency Presentation of Primary Auditory Cortex Following Discrimination Training in Adult Owl Monkeys. *J Neurosci* 1993, 13:87–103.
Adult owl monkeys practicing a tonal discrimination task improved progressively with practice. This improvement correlated with a progressive change in the territory representing the corresponding sound frequencies in the primary auditory cortical field, A1.

34. EDELINE J-M, WEINBERGER NM: Receptive Field Plasticity in the Auditory Cortex During Frequency Discrimination Training: Selective Retuning Independent of Task Difficulty. *Behav Neurosci* 1993, 107:82–103.

35. WEINBERGER NM, JAVID R, LEPAN B: Long-Term Retention of
• Learning-Induced Receptive Field Plasticity in the Auditory Cortex. *Proc Natl Acad Sci USA* 1993, 90:2394–2398.
Conditioned stimulus specific receptive field plasticity in the auditory cortex of adult guinea pigs was found to exhibit long-term retention, evident 1–8 weeks after training. Such learning-induced receptive field modification, established in the awake state, is also expressed under general anesthesia.

36. FREGNAC Y, SHULZ D, THORPE S, BIENENSTOCK E: Cellular Analogs of Visual Cortical Epigenesis. I. Plasticity of Orientation Selectivity. *J Neurosci* 1992, 12:1280–1300.

37. SHULZ D, FREGNAC Y: Cellular Analogs of Visual Cortical Epigenesis. II. Plasticity of Binocular Integration. *J Neurosci* 1992, 12:1301–1318.

38. GILBERT CD, WIESEL TN: Receptive Field Dynamics in Adult Primary Visual Cortex. *Nature* 1992, 356:150–152.

39. CHINO YM, KAAS JH, SMITH EL, LANGSTON AAL, CHENG H: Rapid Reorganization of Cortical Maps in Adult Cats Following Restricted Deafferentation in Retina. *Vision Res* 1992, 32:789–796.

40. PETTET MW, GILBERT CD: Dynamic Changes in Receptive Field Size in Cat Primary Visual Cortex. *Proc Natl Acad Sci USA* 1992, 89:8366–8370.

D Sagi and D Tanne, Department of Neurobiology, Brain Research, The Weizmann Institute of Science, Rehovot 7100, Israel.