VISUAL PERCEPTUAL LEARNING:
A SIGN OF NEURAL PLASTICITY AT EARLY STAGES
OF VISUAL PROCESSING

A. FIORENTINI AND N. BERARDI

Istituto di Neurofisiologia C.N.R., Via S. Zeno 51, 56127 Pisa, Italy

INTRODUCTION

As our ancestors already knew, eventus docet, or, slightly rephrasing it, “prac-
tice makes it better”.

Procedural learning (learning “how”) is a form of implicit (non-declarative)
learning, which involves the acquisition of a new skill through practice. Motor
learning is very familiar: we have all learned to ride a bike, or to skate, or to play
an instrument. Improvement in perceptual skills is perhaps more elusive, although
our ability for instance to discriminate European faces (as opposed to Chinese) or,
for musicians, to discriminate tones is probably the result of perceptual learning.

Almost 15 years ago the existence of a visual perceptual learning with unique
characteristics of selectivity was demonstrated by three groups (3, 4, 12, 13, 26).
Despite the differences in the stimuli and tasks used, all three learning processes
displayed selectivity for the stimulus attributes, such as orientation (12, 26), retinal
location (4, 13) and size (12, 13) suggesting they involved early visual processing.

These seeds fell in the field and took root. Ten years later, there has been a
blooming of papers on perceptual learning and new seeds have probably been sown.
It is time to harvest such a rich crop and store what we have learned from it.

I. Examples of perceptual learning.

Several examples of visual perceptual learning have been reported in the past
and recently. Some of them show an improvement in the detection (e.g. contrast
threshold for detecting oblique gratings, 20) or in the discrimination of visual
stimuli (e.g. the orientation of lines) (29, 30, 36). Some types of hyperacuities, like
stereoacuity (10) and vernier acuity (9, 21, 24), are also improved by practice.
Common to all these visual tasks is the fact that the effects of practice are retained
for a very long time, typically for months or years. Interestingly, however, some
basic functions like grating resolution (5) and the discrimination of gratings dif-
fering for their spatial frequency (13) do not show any improvement with practice,
or is there any appreciable learning for some hyperacuity tasks, like bisection of
the interval between two lines (19) or three-dot alignment (5).

On the contrary, a number of other tasks that require identification or discrimi-
nation of more complex stimuli show a substantial effect of practice, that is
reflected either in the decrease of the discrimination threshold or in a shortening
of the time required for identification. These include global stereopsis (26), discrimination of complex gratings with different luminance profiles (Fig. 1) (12, 13), discrimination of directions of motion (3), motion direction in noise (34), texture segregation (17), global texture identification (2).

II. Selectivity for stimulus parameters and retinal location.

Perceptual learning has a number of properties that are common to all these tasks: it is specific for some parameters of the stimulus, for instance the orientation of lines or gratings (9, 12, 13, 17, 24, 26, 29) (see Fig. 1 B) or the direction of motion (4, 34) and to the retinal location of the stimuli used in the learning procedure (4, 6, 13, 17, 29, 30). For grating stimuli, it is also specific for the spatial frequency (Fig. 1 C) (12, 13). In most cases it is not restricted to the eye used for training, if the training is monocular: learning transfers completely or partially to the untrained eye (3, 13, 29) indicating that the learning process occurs at or centrally to the site where the inputs from the two eyes converge. Texture discrimi-

![Figure 1](image-url)

**Fig. 1.** Perceptual learning in the discrimination of complex gratings with different luminance profiles (first plus third harmonic with relative spatial phases 0 and 90 deg, respectively).

Percentage of correct responses obtained with a forced-choice procedure in the discrimination task. Each point represents percent correct in a block of 10 trials (A and B) or 20 trials (C). A: three subsequent daily sessions, indicating retention of learning in the second and third session. Stimuli: vertical gratings, fundamental frequency 1 c/deg. B: lack of learning transfer from vertical (filled circles) to horizontal gratings (open circles) of the same spatial frequency (1 c/deg). C: lack of learning transfer from vertical gratings of fundamental spatial frequency 3 c/deg (filled squares), to vertical gratings of 6 c/deg (open squares). Learning transfers to vertical gratings of an intermediate spatial frequency (4.5 c/deg, triangles). (For details on the procedure, see Ref. 12, 13).
nation is exceptional, showing little interocular transfer of learning (17, but see 28).

Retinotopy of perceptual learning is very precise in some cases. For instance, for the discrimination of complex gratings with different luminance profiles, learning is restricted to the trained area of 1 deg width, at a retinal eccentricity of 1 deg. No transfer of learning effects is found for gratings located at 2 deg eccentricity, flanking the trained area with no overlap (6). For the discrimination of the direction of motion of random dot patterns there is some degree of transfer if the test stimuli overlap at least partially the trained area (4). In a texture discrimination task, where the location of the training stimuli varies within a retinal quadrant from one trial to the next, the learning has been proved to be restricted to the trained quadrant (17). The same has been found in an orientation discrimination task (30). In this as in other tasks there is no interhemispheric transfer of perceptual learning. An exception however has been reported to the lack of interhemispheric transfer: learning of complex grating discrimination transfers within a narrow strip (1 - 2 deg wide) on either side of the vertical visual field meridian, between areas mirror-symmetric with respect to the meridian. It is likely that in this case transfer of learning is mediated by callosal connections (6).

Information about different stimulus attributes (e.g. luminance and colour; form and motion) is carried by parallel neural channels and processed at partially segregated cortical sites within the primary and secondary visual cortex. Perceptual learning of complex grating discrimination is selective for the chromatic attributes of the stimuli (11). Thus, learning to discriminate gratings having a pure luminance contrast does not transfer to isoluminant gratings defined by a chromatic contrast, and vice versa (Fiorentini and Berardi, in preparation), nor does learning transfer from gratings with a certain chromatic contrast (e.g. a red grating against a green background) to gratings with a different chromatic contrast (blue-green grating against a red background) (Fig. 2 A and B). Interestingly, learning to discriminate complex chromatic gratings is not selective for the orientation of the grating (Fig. 2 C), contrary to what occurs with gratings defined by pure luminance contrast.

The selectivity of visual perceptual learning for basic attributes of the stimuli as well as for their retinal location is suggestive of plastic changes occurring in the visual system at relatively early stages of visual processing, where neurons have relatively small receptive fields and are selective for stimulus attributes such as orientation, size, chromatic properties, direction of motion, etc.

Recently, perceptual learning has been shown to occur in further two visual processes. One is the perception of form from global motion (35), the other is visual search (31, 32). Global motion was obtained by kinematograms consisting of tilted line elements; a rectangular portion of the pattern was shifted to and fro horizontally or vertically, alternating at short time intervals (300 ms). These pattern generate a percept of a rectangle moving horizontally with respect to the background, but for inexperienced subjects it takes many seconds (20, 25 or more) to perceive the rectangle. By repetition of trials, the time required to perceive
Fig. 2. - Percentage of correct responses in the discrimination of gratings of different luminance and chromatic profiles.

Each point represents a block of 40 trials. A: red gratings (650 nm) viewed against a uniform green background (Wratten filter 57), and oriented at -45 deg from vertical. B: blue-green gratings (500 nm) viewed against a red background (Wratten filter 26), same orientation as in A. C: same gratings as in B, but oriented at +45 deg from vertical. Note the lack of transfer of learning effects from A to B, but the transfer from B to C.

form-from-motion decreases considerably, similarly to what occurs for perceiving global stereopsis (26). However, the effects of practice do not transfer from the perception of form-from-motion to global stereopsis (35). Differently from what occurs with global stereopsis, the perception of form-from-motion, once learned, transfers to patterns with elements of different orientation. This is reminiscent of what was found by Ahissar and Hochstein (2) for global identification of texture.

Visual search, i.e. search for a target item in a pattern consisting of a number of different items (distractors), may imply either a parallel, preattentive process or a serial, attentive process. In the first case the time required to detect the target is practically independent of the number of distractors, while in the second case the
time increases with the number of distractors. Steinman (32) and Sireteanu and Rettenbach (31) have provided evidence that a visual search task that is initially "serial" can become "parallel" with practice. In other words, a search that initially requires more time if the number of distractors increases, becomes independent of the number of distractors. Learning is not specific for the item to be searched (for instance a broken circle among full circles) since it transfers to search of completely different items and distractors (for instance a pair of convergent bars among pairs of parallel bars). Thus it seems to imply an improvement in search strategy, rather that an improved perception of the items.

Both these recently reported examples of visual learning differ from many of the examples reported above, in that are not selective for the basic stimulus features present in the patterns. They seem to imply plasticity of the visual system at a level where stimulus generalization is present.

III. Time course of perceptual learning, consolidation and retention.

For some visual tasks a rapid improvement of performance is observed with repetition of trials during a single experimental session and learning effects are partially retained in the second session and totally retained thereafter (see for

![Graph showing performance gain over different sleep conditions](image)

Fig. 3. - *Improvement of performance in a texture discrimination task from the first training session to the subsequent session, run either after a night of normal sleep (NS) or after a comparable night period, but with REM sleep deprivation (REM-) or slow-wave sleep deprivation (SW-).*

SOA is the effective duration of the test stimulus, i.e. the time interval, in ms, between presentation of the texture stimulus and onset of a mask. Performance gain is the decrease from the first to the second experimental session of the SOA interval yielding 80% correct responses. (Redrawn from Ref. 18).
instance Fig. 1 A). In these cases learning is practically complete after a few hundreds of trials (12, 13, 23). For other tasks, there is also a rapid improvement during the first session, but performance continues to improve from one daily session to the next one until a stable optimal level is reached (17, 24). On the whole, thousands of trials are required to complete the learning process (4, 9, 21). In any case, retention is very prolonged: the effects of visual perceptual learning may last for months and even for years.

For texture discrimination a very interesting fact has been reported: improvement in performance from one session to the next only occurs if the two sessions are separated by several hours of rest (17). This suggests that the effects of practice acquired during a training session then trigger a slower process resulting in the consolidation and further improvement of the perceptual skill. Consolidation occurs in the waking state as well as during a normal night sleep, but is strongly dependent on the type of sleep (18). Deprivation of REM sleep prevents any gain in performance on a previously trained task, while non-REM, slow-wave sleep disruption has no detrimental effect (Fig. 3). Thus REM sleep seems to be crucial for transforming the activity-dependent neural change started by visual training into a more stable memory trace, possibly involving a structural modification within the neural network. This confirms the view that sleep is an active process of profound physiological significance as pointed out by Moruzzi (see 22). More specifically, a role for REM sleep in memory consolidation was suggested already in 1970 by Bloch (7) and Pompeiano (25) (for a recent review see ref. 14).

IV. How do we acquire a perceptual skill?

To gain insight into the neural mechanisms involved in perceptual learning requires further characterization of the learning effects. The first point examined is the role of feedback. The paradigm employed in the experiments described above involves the use of feedback, that is, the subject is informed, trial by trial, of the correctness of his/her response.

McKee and Westheimer (21) obtained a practice effect in vernier acuity whether or not they gave their subject feedback. Ball and Sekuler (4) also found that for cardinal directions of movement subjects' accuracy in direction discrimination improved at the same rate whether they were provided of trial by trial feedback or not. For oblique movement directions learning was slower in absence of feedback.

More recently, Fahle and Edelman, (9) and Herzog and Fahle (15), using a vernier displacement discrimination, found that lack of feedback did not prevent learning but only slowed down the process: the slope of the learning curves for observers who received auditory feedback was significantly steeper than for those who did not.

Shiu and Pashler (30) confirmed these observations using a different task, which involved the discrimination of a pair of straight lines differing by 3 deg. They found that subjects learned at the same pace with trial by trial feedback and with block feedback (percentage of correct responses at the end of a block of 44 trials).
This was a “within session” learning, completed in 8-9 blocks. If no feedback at all was provided subjects learned within 3 to 4 days.

A further indication that feedback is not necessary for learning comes from the results of Karni and Sagi (17). In their experiments, subjects performed letter discrimination followed by texture discrimination in the same complex stimulus. Their results show significant learning in the texture task, even though feedback was given only for the letter discrimination.

These results strongly suggest that perceptual learning can proceed in an unsupervised manner, although at a slower pace.

In an attempt to provide a theoretical framework for perceptual learning in vernier hyperacuity, Weiss et al. (37) suggest that a significant component of learning may be based on stimulus driven amplification of neural unit responses which is use-dependent rather than feedback dependent and propose a learning rule, EDL, (Exposure dependent learning) which does not rely on the correlation between a desired output and presynaptic activity to distinguish signal from noise but rather on stimulus-presynaptic activity correlation. Those inputs which are consistently active (i.e. overtake a given threshold) when the stimulus is presented are labelled as signal and undergo amplification. This activity dependent amplification is independent of feedback but is not totally uncorrelated with the activity of the output unit: the synaptic weight of the strongly connected units, whose activity is correlated with the output, would increase more significantly than that of the weakly connected unit, which fails to consistently take part in the activation of the output unit. In other words, when synaptic amplification occurs, it is proportional to the previously acquired synaptic strength. This assumption of Weiss et al. (37) adds a Hebbian element to the learning rule they propose as a basis for unsupervised learning, which is then classified as “non interactive Hebbian rule”. A stronger accent on the requisite of correlation in pre- and post-synaptic activity in order to increase the strength of a neural connection (Hebbian rule) is given in the paper by Vaina et al. (34). The conclusions of these authors are, nonetheless, in favour of a model of unsupervised learning process, independent on external feedback, in which “noisy” neurones, that do not contribute to the perceptual judgement are progressively dropped out from the computation: this model learns to ignore noise. It is interesting that in Vaina et al. psychophysical experiments, subjects learned only if their initial performance was above chance level. If feedback is not essential to perception improvement, does the mere repeated exposition to the stimuli lead to learning? A number of experiments have concluded that this is not the case. Practice effects are not determined solely by activity in stimulus driven mechanisms but also by high level attentional mechanisms, which probably control, with a top-down effect, changes taking place at early visual processing levels.

Ahissar and Hochstein (2) designed an elegant experiment, in which two independent tasks could be performed on the same set of visual stimuli, with very similar degree of difficulty. The stimulus was a briefly presented array of short line segments. Segments were arranged in a vertical or horizontal array and one of the
segments had a different orientation from the others. Observers either identified the orientation of the array (global task) or detected the presence of the odd man out (local task). As a consequence, after practice on one task, the experimenters were able to test the presence of any improvement in the other task, while keeping the retinal stimuli constant. The result was very clear: learning did not take place for the stimulus attributes that were not attended to.

Shiu and Pashler (30) obtained very similar results with observers practicing in discriminating the brightness of two lines which also differed for their orientation: orientation discrimination was not improved at the end of the brightness task.

The suggestion which comes out is that attention controls performance of neural networks which operate at low, stimulus related processing levels. Attributes which are processed by different networks do not share the benefits of being highlighted by attention. Rather, practicing on one task while using a set of neural elements might even lead to unlearning in a task which involved a different set of neural elements (33).

It is interesting that a "modulatory input", which is provided by neurones excited by motivationally significant stimuli, is incorporated in the model of Weiss et al. (37).

These conclusions lead us to the ultimate question: the cellular mechanisms underlying perceptual learning.

No study in the literature has explicitly addressed this question. A recent paper by Ahissar et al. (1) presents the results of a very ingenious and elegant study on the neural correlates of learning (conditioning) in which the relevance of the behavioural context (attention) was also investigated. They studied the correlation between firing times of pairs of neurones in the auditory cortex. After assessing the actual correlation exhibited by each pair of neurones, they applied a conditioning paradigm in which an auditory stimulus capable of eliciting or suppressing activity in one neuron of the pair (US stimulus) was delivered triggered by the occurrence of one spike in the other neuron of the pair, thus forcing the former to be positively or negatively correlated to the latter. The conditioning paradigm (duration 70 to 850 sec) was applied either to monkeys performing an auditory task in which they had to detect the US stimulus (behaving monkeys) or to monkeys not attending to the stimuli (control monkeys). The results indicated that the functional connections between auditory neurones, as evaluated by the degree of cross correlation, were strengthened in the case of positive correlation and weakened in the case of negative correlation (functional plasticity), but only in behaving monkeys. The effects lasted for a few minutes after the end of the conditioning. Almost no effect was seen in control monkeys. This suggests both that experience determines rapidly changes in synaptic efficacy depending on the level of contingency and that this functional plasticity is strongly affected by behavioural factors.

Poggio (23) and Poggio et al. (24) proposed that the perceptual modules within which practice determines use dependent, feedback independent amplification of unit responses, might either be hard-wired from birth or newly synthesized in a task driven fashion. Particularly for the examples of fast perceptual learning here
reviewed, the idea that in our brain new task specific modules are created is appealing; the acquisition of new ability might therefore correspond to the acquisition of new committed local networks. The increase of correlated activity within one area could be a sign of this process.

But which areas should we be looking to, in order to deeper understand perceptual learning?

The characteristic selectivity for the stimulus parameters of the learning processes here reviewed points to visual areas where neurones have small receptive fields and exhibit selectivity for the stimulus attributes, that is, areas at a low hierarchical level. The various authors stops this far, “early visual processing”, without making any further suggestion. Only Karni and Sagi (17), on the basis of the lack of interocular transfer found in their learning process venture to propose area V1 as a possible neural locus for the underlying plastic changes. Information on the areas involved in a given function generally comes from two types of experiments: studies in patients or animals with specific and restricted lesions or, more recently, studies with imaging techniques such as PET and fMRI. Neither type is particularly present in the literature of perceptual learning. Schiller and Lee (27), studying the performance of monkeys with V4 lesions, suggest that this area is more important for the generalization of discrimination learning over the visual field (overtaking strict retinotopy) than for discrimination learning per se. Since this is more a problem of generalization than of learning, one would conclude that learning had taken place at yet lower visual areas. Courtney et al., (8) present the most exciting and promising abstract in the quest for the areas involved in perceptual learning: they find that extrastriate visual areas activity is correlated with visual priming and, more to our point, that early visual areas undergo long lasting changes of activation while the subject learns the task employed by Karni and Sagi (17). How early this areas are is only a matter of time to know.

SUMMARY

Examples of perceptual learning in various visual tasks are briefly reviewed. In spite of the variety of the tasks and stimuli, in most of these examples the effects of learning are specific for stimulus parameters and retinal location and transfer interocularly. Enduring practice effects can be acquired within a single experimental session and/or progressively from one session to the next one, often continuing to improve until thousands of trials have been performed. The consolidation of learning effects from one session to the next one may occur in the waking state as well as during a normal night sleep, but is strongly dependent on the type of sleep. Improvement in performance does not require that the subject is informed of the correctness of his/her responses, but needs attention to the task: learning does not take place for the stimulus attributes that are not attended to. All this suggests that visual perceptual learning involves plastic changes at early neural processing levels, which are dependent for their induction and consolidation on the general behavioural state of the subject, such as attentiveness and type of sleep.
REFERENCES


