

Three remarks on perceptual learning

J. D. MOLLON¹ and MARINA V. DANILOVA²

¹*Department of Experimental Psychology, University of Cambridge, Downing St., Cambridge, CB2 3EB, UK*

²*Visual Physiology Laboratory, I. P. Pavlov Institute of Physiology, Academy of Sciences of Russia, Nab. Makarova 6, St Petersburg, 199034, Russia*

Received 23 July 1995; accepted 18 October 1995

Abstract—This essay makes three points. (1) From the failure of perceptual learning to transfer when stimulus parameters are changed, it cannot necessarily be concluded that the site of learning is distal: rather, the learning may be central and the specificity may lie in what is learnt. (2) Mere exposure to a stimulus may not be sufficient for learning: even in the absence of explicit feedback, other sources of information are often available in a perceptual learning task. (3) Procedural learning may sometimes be apparent only after a delay: this phenomenon has a long history and is known as reminiscence.

Our limited purposes in this Note are to distinguish two explanations of the specificity of perceptual learning, to consider what sources of information are available to the learner, and to recover from the classical literature some papers neglected in recent discussions.

1. ALTERNATIVE EXPLANATIONS OF THE SPECIFICITY OF PERCEPTUAL LEARNING

Many examples of perceptual learning have proved to be very specific, in that little transfer of the learning is observed when the stimulus conditions are changed. For example, subjects improve over thousands of trials in a vernier-acuity task but show little transfer when the target is changed in its orientation or position or is presented to the other eye (Poggio *et al.*, 1992; Fahle and Edelman, 1993; Fahle, 1994). A popular interpretation of such results has been that the site of learning is very distal in the visual system, that specialized processors or ‘modules’ are developed at a relatively early level of analysis or that the tuning of existing channels is increased. Thus Fahle (1994) writes ‘The orientation specificity we found requires that the neurons that learn are orientation specific . . . the fact that learning is mostly eye specific suggests that the neurons that learn are mostly monocular . . . the position specificity of learning indicates that the underlying neuronal processes occur in a cortical area where position invariance has not yet been achieved. These results suggest Area VI as the most probable candidate for learning of visual hyperacuity.’ Similarly, Shui

and Pashler (1992), having found the learning of orientation discrimination to be specific to retinal position, concluded that the learning occurs 'in visual areas that are still retinotopically organized.' See also Ahissar and Hochstein (1993) and Sagi and Tanne (1994).

There is an obvious and less exciting alternative hypothesis that needs to be ruled out in each such case: the *site* of the learning may in fact be central and what is specific may be *what* is learnt. Indirectly, via the central signals to which he has access, the subject may be learning about optical features of his retinal image; about the local topography of his receptor mosaic; and about the specific wiring of individual neurons within his visual pathways. We briefly consider these three factors in turn:

(i) As Le Grand nicely put it, the retinal image is the one optical image that was never intended to be seen: the image is degraded by several optical aberrations, some common to all eyes, others peculiar to particular eyes and indeed to particular retinal regions. The sources of optical degradation include diffraction, spherical aberration, chromatic aberration, astigmatism, and the blur produced by eye movements. So the first thing that the subject has effectively to learn is the point-spread function for the retinal region and the viewing distance—the accommodative state—that the experimenter has chosen for the task.

(ii) Secondly, the matrix of photoreceptors is not perfectly regular: the hexagonal array of foveal cones is subject to local jittering and to fault lines. This disorder is apparent in photomicrographs of the retina (Polyak, 1949; Borwein *et al.*, 1980), but is also suggested by the *in vivo* analyses of D. Williams, which are based on the moiré fringes seen when interference gratings of high spatial frequency are imaged on the retina (Williams, 1988). In a hyperacuity task, by definition, the subject discriminates distances that are less than the diameter of a cone, but the image of a narrow target is known to extend across several cones. When a vernier target is first flashed on a particular retinal region, the naive subject cannot be expected to know exactly which ratios of cone absorptions correspond to rightward displacements of the vernier and which to leftward. He must slowly learn to interpret the patterns of relative absorptions that correspond to different members of the set of stimuli that the experimenter has chosen. Effectively, he is learning, perhaps for the first time, about the detailed arrangement of cones in a very specific region of one eye.

(iii) Thirdly, there is necessarily a biological variability in the connections between a local sub-array of receptors and the several parallel post-receptoral channels that draw their inputs from that sub-array. To achieve the maximal performance in the task, the subject has effectively to learn about these connections for a specific region of the visual field.

We assume, of course, that the subject has no direct access to his point-spread function, to his retinal matrix or to the wiring of channels in his early visual system. What in practice he must learn is which tiny subset of central channels is most useful for sorting the set of discriminanda that the experimenter has settled upon; and how variations in the signals of these channels map on to the external discriminanda. If now the target is transferred to a different retinal region or is presented to the other eye or is changed in orientation, the optical properties of the image, the local arrangement

of the receptor array, and the wiring of feature detectors will be different. The subject must learn to interpret a new set of neural signals¹.

In an experiment intended to discover what is learnt in a vernier acuity task, Saarinen and Levi (1995) have shown that learning on the primary task is accompanied by a narrowing of the range of mask orientations that interfere with the hyperacuity task. They hypothesize that learning may depend on a change in the orientational tuning of cells in the early visual system, but an alternative possibility is that there has been a change in the subset of signals on which the subject relies; and this explanation is close to the alternative that Saarinen and Levi themselves allow in their final paragraph.

In summary, specificity of learning for particular stimulus conditions does not *in itself* imply that the site of learning is distal. The learning may occur at a central site, and what the subject may be learning about are the local idiosyncracies of his retinal image, of his receptor mosaic, and of the wiring of his visual system. Of course, other kinds of experiments may bear on the site of the plasticity (Fahle, 1994), but these must be evaluated separately and our concern here is only with the logic of the argument from specificity.

2. WHAT SOURCES OF INFORMATION DOES THE SUBJECT USE TO CALIBRATE HIS IMAGE? IS PURE EXPOSURE SUFFICIENT FOR PERCEPTUAL LEARNING?

Although some studies have shown more rapid learning when feedback is given (Fahle and Edelman, 1993), perceptual learning occurs even if the observer is not given explicit feedback (McKee and Westheimer, 1978). The latter observation has probably reinforced the hypothesis that the learning depends on the formation of specialised modules early in the visual system. The subject does better, it is supposed, simply because pure exposure, often repeated, leads to the formation of a module specialised for analysing the arbitrary target set that the experimenter has devised².

The hypothesis is an attractive one, but whether or not the site of plasticity is thought to be distal or central, we should ask whether any information is in fact available to a subject during a perceptual learning experiment without feedback. If the exposures are longer than, say, 200 ms, information may be provided by gross eye-movements: if a subject makes saccades from one part of a target to another, an error signal is provided by the discrepancy between the landing position and the expected position. A classical example of perceptual learning is the complete or partial abolition of the Müller-Lyer illusion after hundreds of presentations. This increase in accuracy occurs when free eye movements are allowed but not if the subject is asked to fixate (Lewis, 1908).

Even if the exposure is relatively short, the fine-scale movements of the eye—tremor, drift and microsaccades—may supply feedback in hyperacuity tasks. Many years before the present excitement at the learning capacity of artificial nets, J. R. Platt explicitly asked how a neural net connected to a disordered matrix of photoreceptors could learn to recognize straight lines (Platt, 1958, 1960). He pointed out that a straight line has the special property of 'self-congruence under displacement': if a small linear eye movement is made along the line, the image of the line is displaced

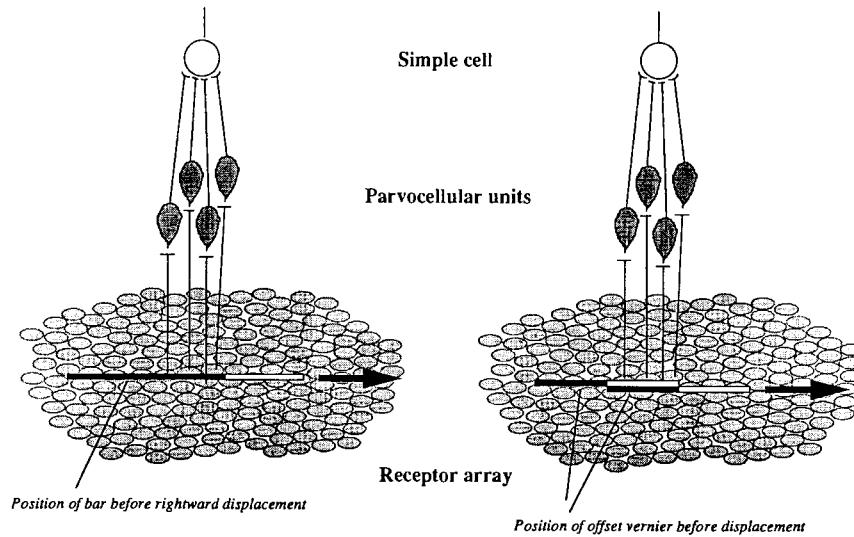


Figure 1. An interpretation of Platt's theory of how the retinal metric is calibrated. In the diagram, the black bars represent the target before the eye movement—a straight line in the left-hand example and an offset vernier in the right-hand case. In the left-hand example, the straight line is displaced along its own length as a result of a small lateral eye movement (the white bar represents its new position). Via midget bipolars, midget ganglion cells and parvocellular LGN units, a simple cell draws inputs from a region in the centre of the line's path. The input to the simple cell remains constant before and after the movement, and this is an indication that a line (or edge) is self-congruent under displacement. In the example to the right, the displaced line is not rectilinear in this local region and the input to the simple cell does not remain stable during a linear eye movement. But does the calibration process require a reafferent signal from the eye-movement control system and how is the trajectory of the eye movement known with sufficient precision? Platt is not explicit on this issue. However, he points out that a perfect arc exhibits self-congruence under displacement when an eye movement is made along a curved path. He thus implies that the calibration process requires either afferent or reafferent information about eye movements, since this information is needed to distinguish between the self-congruence of a straight line and that of an arc. Yet Platt himself opens his discussion with the analogy of lens grinding, where the relative movement of two roughly curved surfaces can result in a spherical surface of much higher precision than either. Now, in the process of lens grinding, it is essential that the two surfaces move independently: there is no coupling, no communication, between them. If we assume that the basic calibration of the visual metric is strictly in terms of locally straight segments of contour and that fine eye movements are predominantly linear, then we can envisage that learning might occur through the accumulation of instances of displacement with self-congruence for many different directions—and different simple cells. This hypothesis differs from Platt's in supposing no communication between the retinal afferent system and the eye-movement control system.

along itself (Fig. 1). If the line is regular and is appropriately aligned, there will be a set of photoreceptors—those on which the image of the line falls throughout the movement—whose signal remains invariant before and after the eye movement. All that the neural net is required to do is to recognize invariance. It does not need knowledge of the exact positions of the receptors; and the actual retinal image of the line does not need to be straight. Platt does not himself refer to the work of Hubel and Wiesel (1959), which was not published when he was developing his hypothesis, but it is instructive that rectilinear lines and edges are the primitives that the primary visual cortex appears to extract. The property of self-congruence under displacement means that the visual system would have an obvious means of calibrating cells that extracted these primitives³. Support for Platt's hypothesis might be seen in the finding of Bennett and Westheimer (1991) that learning did not occur in a task in which the observer was required to detect the misalignment of three points: this hyperacuity task seems closely to resemble the vernier task, where clear learning does occur, but differs from it in that only very rarely will an eye movement be of the correct direction and magnitude to produce a self-congruent displacement of the dots.

If the experimenter adopts the Method of Constant Stimuli (as is traditional, for example, in the case of vernier-acuity tasks), then there is still at least one further source of information available to the subject and that is the range of stimuli used in the experiment. Owing to local features of the retinal mosaic, all or most of a symmetrical set of vernier targets might initially appear displaced in the same direction (Morgan 1992), but the knowledge (or the implicit assumption) that the experimental set is centred on the point of collinearity may allow the observer to reinterpret the sensory inputs. The correction might be at a cognitive level, but alternatively information about the statistical distribution of stimuli might be used by more automatic, distal mechanisms—as in the hypothesis of Andrews (1964), to which Morgan (1992) has drawn attention. According to Andrews, the calibration of visual space depends on a process similar to the *normalization* postulated for curvature and movement by Gibson (1937) or the *adaptation level* mechanism postulated for colour by Helson. Initially, Andrews' receptor matrix is coupled only topologically to the central metric, but the mapping is adjusted over time so that in the central representation (a) the average curvature at any point is zero and (b) the average spacing of contours is equal throughout the visual field. If there is no bias in the sense and amplitude of curvatures in the observer's world, and if the external distribution of contours is random when averaged over time and space, then the central metric will come to be topographically accurate. Andrews does not himself deal specifically with the detection of vernier offsets. In intercourse with typical visual environments, his two processes of normalization might yield a general metric adequate for the offsets we commonly need to discriminate; but it is not obvious how his two normalizations (*a* and *b* above) would be of use in the reduced conditions of a vernier task. One solution would be to suppose that the vernier task was solved by analyzers also intended for discrimination of curvature. Another would be to postulate a third process of normalization: one that ensures (c) that in the central representation the average local contour or edge is continuous, that is, offsets of opposite sense are equally common. It is a plausible, if unproven, assumption that offsets of opposite sign

are equally frequent in visual scenes (and perhaps are statistically rare). There is one last assumption that we might add to those of Andrews, and this would be that the normalization processes are carried out in parallel for different spatial frequencies. Usually, the metrics of the several representations would be in agreement, but illusions could arise when the local statistics of the stimulus were abnormal in a particular spatial frequency band.

The foregoing considerations suggest that a good test of whether pure exposure improves performance would be (a) to reduce stimulus duration to units of milliseconds and (b) to use an adaptive staircase procedure that allows drift in the point of subjective equality. The former would deprive the subject of the information available from fine-scale eye movements, the latter would deprive him of the information that comes from a fixed or symmetrical stimulus distribution.

3. REMINISCENCE: IMPROVEMENT IN PERFORMANCE AFTER A DELAY

In an experiment on backward masking of texture patterns, Karni and Sagi (1993) required the subject to report the overall orientation of a sub-array of three line-elements that differed in orientation from the background elements. The target array was followed by a pattern mask and the dependent variable was the stimulus-onset-asynchrony giving 80% correct discrimination. Karni and Sagi found that performance was stable within a session and that long-term improvements revealed themselves only at intervals of eight hours or more after the end of the last session. Similarly, in their study of orientation discrimination, Schoups *et al.* (1995) found that there was very little improvement across the 16 blocks of trials of an individual session, whereas there was clear improvement between sessions conducted on successive days.

Karni and Sagi (1993) write 'perhaps for the first time, our results show that not all human learning is concurrent with practice'. In fact, learning phenomena with this behaviour are well known in the classical literature on the acquisition of skills, and there pass under the name of *reminiscence* (McGeoch and Irion, 1952; Woodworth and Schlosberg, 1954). Sutherland (1989) gives the following definition of 'reminiscence': *The phenomenon that after a task has been partially learnt, performance on it may be better when tested some time after learning has occurred than when tested immediately afterward.* Travis (1937a), for example, used the pursuit rotor task, in which the subject must maintain maximum contact between a hinged, weighted stylus and a metal contact embedded in a disc rotating at one revolution per second. Travis found little improvement during the actual training sessions lasting 6 min, but after a rest period of either 3 or 4 days the subjects exhibited a much better level of performance. In subsequent experiments with the pursuit oscillator or the pursuit rotor, the optimum length of the rest period was found to be shorter than that observed by Karni and Sagi—of the order of 10–60 min (Travis, 1937b; Ammons, 1947; Kimble and Horenstein, 1948)—although the optimum length is known to depend on the extent of preceding practice and on the number of trials used in the estimate of subsequent improvement (Bell, 1942; Melton, 1941; Norris, 1953). Learning on such tasks, like the learning studied by Karni and Sagi, may survive unimpaired over an interval of the

order of two years. And the explanation adopted by Karni and Sagi—consolidation over time of memory traces—is one of the standard theories of reminiscence in the classical literature (McGeoch and Irion, 1952; Eysenck, 1965).

Although an acceleration of learning on the Müller-Lyer figure has been observed when the interval between sessions was increased from one day to two (Lewis, 1908), it is true that most of the explicit demonstrations of reminiscence have been in the field of perceptual-motor learning rather than purely perceptual learning. However, Karni and Sagi's visual task falls in the same general category of procedural learning—the acquisition of skill—and it may prove both economical and illuminating to consider their result within the more general context of reminiscence effects. At the very least, their finding is hardly without antecedents in the field of human learning.

'... we learn to swim during the winter and to skate during the summer.' (Wm. James *Principles* vol. I, p. 110)

Acknowledgement

Preparation of this paper was assisted by a Physiological Society travel grant to M. V. Danilova. We are grateful to H. B. Barlow, M. J. Morgan, and B. C. Regan for comments on the text.

NOTES

1. The subject will, of course, learn some general features of the task—such as the timing of the stimuli, the responses required, and, where they are offered, the meaning of feedback signals. The learning of these features probably contributes to the early 'fast' phase of learning and to any partial transfer that occurs when stimulus conditions are changed.

2. In an orientation discrimination task, Shiu and Pashler (1992) found that some improvement occurred without feedback but did require attention to the relevant stimulus dimension: extended judgements of the relative brightness of two lines did not enhance the discrimination of their orientation in a transfer test.

3. Once the system is calibrated, of course, Platt's hypothesis does not require eye movements for performing the vernier task. This point was not fully grasped by early commentators such as Riggs (1960).

REFERENCES

- Ahissar, M. and Hochstein, S. (1993). Attentional control of early perceptual learning. *Proc. of the National Academy of Sciences of the United States of America* **90**(12), 5718–5722.
- Ammons, R. B. (1947). Acquisition of motor skill: II Rotary pursuit performance before and after a single rest. *J. Exp. Psychol.* **37**, 393–411.
- Andrews, D. P. (1964). Error-correcting perceptual mechanisms. *Q. J. Exp. Psychol.* **16**, 104–115.
- Bell, H. M. (1942). Rest pauses in motor learning as related to Snoddy's hypothesis of mental growth. *Psychol. Monographs* **54** (Whole No. 243).
- Bennett, R. G. and Westheimer, G. (1991). The effect of training on visual alignment discrimination and grating resolution. *Perc. Psych.* **49**, 541–546.
- Borwein, B., Borwein, D., Medeiros, J. and McGowan, J. W. (1980). The ultrastructure of monkey foveal photoreceptors, with special reference to the structure, shape, size, and spacing of the foveal cones. *Am. J. Anat.* **159**, 125–146.

- Eysenck, H. J. (1965). A three-factor theory of reminiscence. *British J. Psychol.* **56**, 163–181.
- Fahle, M. (1994). Human pattern recognition: parallel processing and perceptual learning. *Perception* **23**, 411–427.
- Fahle, M. and Edelman, S. (1993). Long-term learning in vernier acuity: Effects of stimulus orientation, range and of feedback. *Vision Res.* **33**, 397–412.
- Gibson, J. J. (1937). Adaptation with negative after-effect. *Psychol. Rev.* **44**, 222–244.
- Hubel, D. H. and Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *J. Physiol.* **148**, 574–591.
- Karni, A. and Sagi, D. (1993). The time course of learning a perceptual skill. *Nature* **365**, 250–252.
- Kimble, G. A. and Horenstein, B. R. (1948). Reminiscence in motor learning as a function of length of interpolated rest. *J. Exp. Psychol.* **38**, 239–244.
- Lewis, E. O. (1908). The effect of practice on the perception of the Müller-Lyer illusion. *British J. Psychol.* **2**, 294–306.
- McGeoch, J. A. and Irion, A. L. (1952). *The Psychology of Human Learning*. Longmans, Green and Co., New York.
- McKee, S. and Westheimer, G. (1978). Improvement in vernier acuity with practice. *Percept. Psychophys.* **24**, 258–262.
- Melton, A. W. (1941). The effect of rest pauses on the acquisition of the pursuitmeter habit. *Psychological Bulletin* **38**, 719.
- Morgan, M. J. (1992). Hyperacuity of those in the know. *Current Biology* **2**, 481–482.
- Norris, E. B. (1953). Performance of a motor task as a function of interpolation of varying lengths of rest at different points in acquisition. *J. Exp. Psychol.* **45**, 260–264.
- Platt, J. R. (1958). Functional geometry and the determination of pattern in mosaic receptors. In: *Symposium on Information Theory in Biology*. H. P. Yockey, R. L. Platzman and Quastler H. (Eds). Pergamon, London, pp. 371–398.
- Platt, J. R. (1960). How we see straight lines. *Scientific American* **202**(6), 121–129.
- Poggio, T., Fahle, M. and Edelman, S. (1992). Fast perceptual learning in visual hyperacuity. *Science* **256**, 1018–1021.
- Polyak, S. (1949). Retinal structure and colour vision. *Documenta Ophthalmologica* **3**, 24–46.
- Riggs, L. (1960). Correspondence. *Scientific American* **203**(5), 16.
- Saarinen, J. and Levi, M. D. (1995). Perceptual learning in vernier acuity: what is learned? *Vision Res.* **35**, 519–527.
- Sagi, D. and Tanne, D. (1994). Perceptual learning: learning to see. *Current Biology* **4**, 195–199.
- Schoups, A. A., Vogels, R. and Orban, G. (1995). Human perceptual learning in identifying the oblique orientation: retinotopy, orientation specificity and monocularly. *J. Physiol.* **483**.3, 797–810.
- Shiu, L.-P. and Pashler, H. (1992). Improvement in line orientation discrimination is retinally local but dependent on cognitive set. *Percept. Psychophys.* **52**, 582–594.
- Sutherland, S. (1989). *MacMillan Dictionary of Psychology*. MacMillan, London.
- Travis, R. C. (1937a). Practice and rest periods in motor learning. *J. Psychol.* **3**, 183–187.
- Travis, R. C. (1937b). The effect of the length of rest period on motor learning. *J. Psychol.* **3**, 189–194.
- Williams, D. R. (1988). Topography of the foveal cone mosaic in the living human eye. *Vision Res.* **28**, 433–454.
- Woodworth, R. A. and Schlosberg, H. (1954). *Experimental Psychology*. Methuen, London.