The reverse hierarchy theory of visual perceptual learning

Merav Ahissar[1] and Shaul Hochstein[2]

[1]Department of Psychology and Interdisciplinary Center for Neural Computation, Hebrew University, Jerusalem, 91905, Israel
[2]Department of Neurobiology and Interdisciplinary Center for Neural Computation, Hebrew University, Jerusalem, 91905, Israel

Perceptual learning can be defined as practice-induced improvement in the ability to perform specific perceptual tasks. We previously proposed the Reverse Hierarchy Theory as a unifying concept that links behavioral findings of visual learning with physiological and anatomical data. Essentially, it asserts that learning is a top-down guided process, which begins at high-level areas of the visual system, and when these do not suffice, progresses backwards to the input levels, which have a better signal-to-noise ratio. This simple concept has proved powerful in explaining a broad range of findings, including seemingly contradicting data. We now extend this concept to describe the dynamics of skill acquisition and interpret recent behavioral and electrophysiological findings.

Throughout life our sensory receptors are continuously bombarded by stimuli. This activation not only induces perception, it also modifies our representation mechanisms, thereby affecting all subsequent perception. Recent evidence shows that a large degree of perceptual plasticity is retained in adulthood, with long-term manifestations including adaptation, priming and perceptual learning. Currently, these different phenomena are defined by their behavioral characteristics and the manner in which they are induced, rather than by their respective underlying neural mechanisms.

This article focuses on perceptual learning, defined as practice-induced improvement in the ability to perform specific perceptual tasks (see [1] for a classic review and [2] for recent overviews). We shall argue that what typically limits naïve performance is the *accessibility* of task-relevant information rather than the absence of such information within neuronal representations [3]. We shall present the reverse hierarchy theory (RHT) of perceptual learning asserting (i) that perceptual improvement largely stems from a gradual top-down-guided increase in usability of first high- then lower-level task-relevant information, and (ii) that this process is subserved by a cascade of top-to-bottom level modifications that enhance task-relevant, and prune irrelevant, information (see [4] for a computational model applying a similar concept).

The relations between plasticity processes during substantial practice and those dominating the first few exposures, are not well understood. Only the first are

typically referred to as perceptual learning, whereas the latter are termed priming. Recent evidence suggests that, at least when governed by top-down control, single exposures (priming) can induce strong and long-lasting effects that clearly change our perception (see Box 1), suggesting that these might be the initial processes of perceptual learning, as described below (The Eureka effect).

Box 1. The Eureka effect

The picture (Figure I) generally appears simply as a set of gray and black regions, without further meaning; the object represented is hard to categorize without further cues. Following considerable inspection, (or a single look at the 'clue' in Figure 1 overleaf), the puzzle is solved, and, without further practice, observers directly perceive a bearded figure. This effect is rapid, strong and long-lasting, suggesting that significant top-down control determines our conscious perception. Ahissar and Hochstein [21] found that a single long exposure to a 'pop-out' stimulus enabled learning of a very difficult detection task, based on brief and strongly masked presentation of similar stimuli – a task that was almost never learned without the Eureka enabling experience. Thus, it appears that similar top-down control or guidance mechanisms influence both perceptual learning and conscious perception. An important difference between these effects is that in the experiment, following the single easy-case 'Eureka' exposure, hard-case perceptual learning was enabled, but still required; no such training is needed for Figure I here.
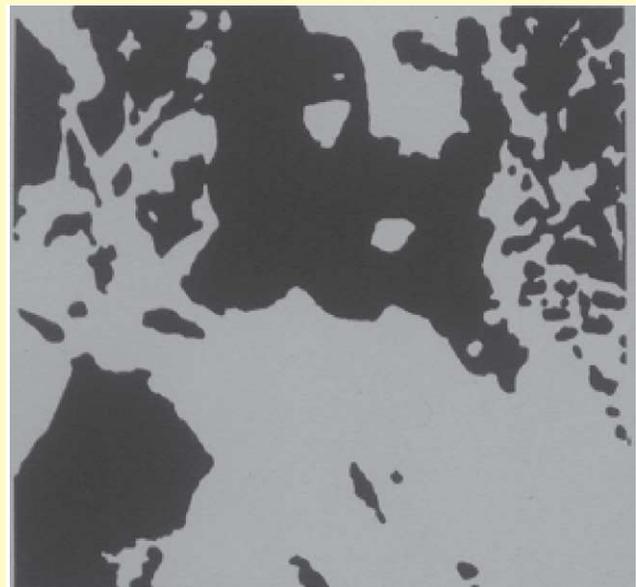

Figure I.

Corresponding author: Merav Ahissar (msmerava@pluto.huji.ac.il).

Figure 1. A hint for 'understanding' the image in Box 1. A few lines draw attention to those aspects that are essential for perceiving a bearded figure. Now return to Box 1 on the preceding page to see how this 'clue' has affected your perceptual system, perhaps forever!

A third phenomenon of perceptual plasticity is adaptation. This seems to be a more basic mechanism that is not specific to the nervous system and is characteristic of all biological systems. It differs from perceptual learning in being induced by exposure to stimuli rather than by task-specific practice. Thus, it is essentially a bottom-up process, in which internal representations are modified in response to the current distribution of external stimuli [5,6]. Consequently, responses to unvarying stimuli are reduced, inducing increased sensitivity to changes, or novelty detection. Because perceptual training involves exposure to stimuli and consequently adaptation, it is difficult to study perceptual learning without inducing adaptation processes, whose impact on task performance is hard to discern (see discussion in [7] and an experimental example of bias induced by adaptation in [8]).

Psycho-anatomy logic and the Reverse Hierarchy Theory
The term psycho-anatomy, coined by Julesz more than 30 years ago [9], implies that we can deduce from behavioral findings information regarding the underlying anatomical structures. In the visual modality, our knowledge of basic representations is relatively broad, based mainly on single unit receptive field characteristics, and recently on corroborations from fMRI studies, [10–12]. Particularly well understood is the representation of oriented light or dark bars and edges. In the primary visual cortex, V1, single neurons are selective for orientation and retinal position with relatively narrow tuning curves in these domains (the 'Simple' cells of Hubel and Wiesel, [13]). Both spatial and orientation tuning curves are broader for neurons at higher visual areas along the visual hierarchy.

This basic observation can be used to deduce the site of neuronal modifications from learning generalization. Namely, if we train subjects on a perceptual task using a set of stimuli whose orientation and retinal position are fixed, then changing the position or orientation of the stimulus would lead to activation of a non-overlapping population of neurons at lower-level areas. Thus, if the major bulk of plasticity underlying behavioral improvement occurred at low-levels, improvement would not transfer to these new stimulus conditions and subject performance would be degraded towards initial levels, requiring a process of re-learning. On the other hand, if learning resulted from high-level modifications, it would largely transfer to novel positions and orientations. These possibilities are schematically illustrated in Figure 2 for Simple and Complex V1 neurons [13], and higher-area neurons (inferotemporal cortex, IT; [14]).

This psycho-anatomy logic has been used in many psychophysical studies to deduce the underlying site of plasticity from the degree of learning specificity to spatial dimensions such as the trained eye [15], retinal position and orientation (e.g. [16–19]). However, parallel studies produced contradictory findings regarding learning specificity even when experiments seemed rather similar. For example, Karni and Sagi [15] found remarkable specificity, including to the trained eye, whereas Schoups *et al.* [20], using a very similar texture discrimination task, found complete generalization across eyes.

We decided to study directly whether this variability itself follows systematic rules. The task we used was detecting the presence of an oddly oriented bar in an array of homogenously oriented distractor bars [17–19,21,22], as illustrated in Figure 3. In most of the stimuli that we used, the orientation of the target greatly deviated from that of the distracting bars, yielding effortless detection, accompanied by the notion that the odd element 'pops out' ([23]; see review in [24]). That is, with long exposures (>250 ms), this task is trivial to begin with, and performance (reaction time) is independent of the number of distractor elements in the array. We made it difficult by using brief exposures followed by a masking stimulus (Figure 3). Under these conditions, task difficulty could be controlled in several ways. Most frequently we manipulated the functional stimulus-processing time by varying the interval between stimulus and mask (stimulus-to-mask onset asynchrony; SOA). With practice, the minimal SOA required for threshold detection (e.g. achieving 80% correct) was substantially reduced [17].

The orientation specificity of this improvement indeed followed a consistent pattern: when stimulus detection was made difficult by a brief processing time (short SOA), small target/distractor orientation difference, increased target position uncertainty, or a target presented farther from fixation, improvement was slower and orientation specific. By contrast, learning of easy cases was general and transferred across orientations. This pattern was found both across groups training under easy or difficult conditions, respectively, and across easy and difficult conditions within the same individuals. Thus, practicing orientation detection with interleaved SOAs, observers showed generalization for long (easy) SOAs and specificity for short (difficult) SOAs. Moreover, when learning was orientation specific, it was also position-specific, consistent with the linkage of orientation and position specificity
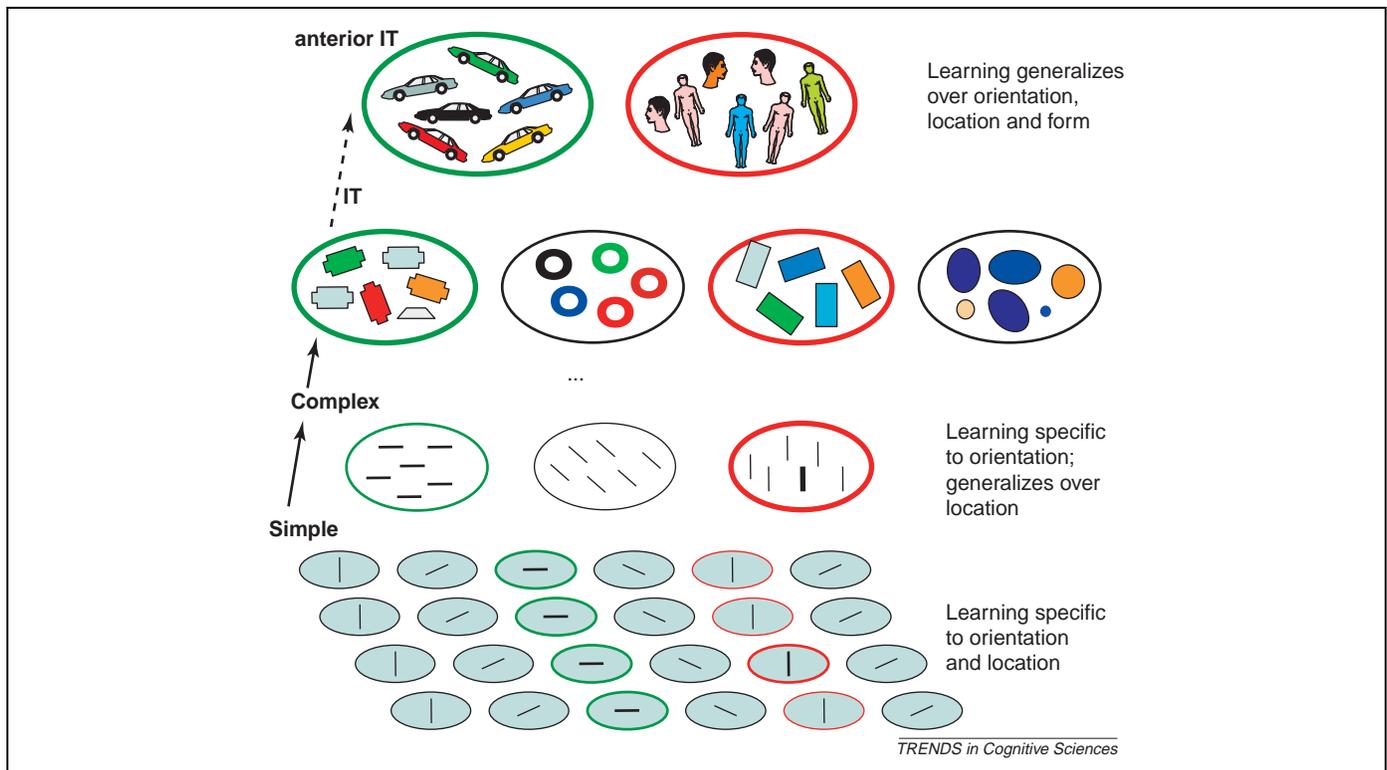
Figure 2. The psycho-anatomy logic and the visual hierarchy: specificity versus transfer in perceptual learning. A hierarchy of cortical areas progressively processes visual information, representing first specific details of color, orientation, motion, 2D position and stereoscopic disparity. Higher level representations generalize over these parameters and specialize in objects, categories and concepts. Schematic drawings of these levels are shown above. First, striate cortex 'Simple' and 'Complex' receptive fields respond to specific orientations, (colors and line widths or spatial frequencies), with Simple cells being more specific for stimulus location (and perhaps less so for spatial frequency) than Complex cells. Following intermediate processing levels, different regions of inferotemporal (IT) cortex represent basic shapes, and object categories. These units generalize over the basic dimensions of the stimulus. If perceptual task training is to affect performance, it must modify representations at one or more of these levels. If modification is at very specific lower level units, learning effects will not transfer to performance of tasks with stimuli of different color, orientation, etc, because the neurons responding to these stimuli were not affected by training. On the other hand, if training-induced modification is at high levels, in broadly-tuned units, then learning effects will transfer to new stimulus conditions. Thus, testing the degree of specificity or transfer of learning effects can inform us of the level of the training-induced neural modification.
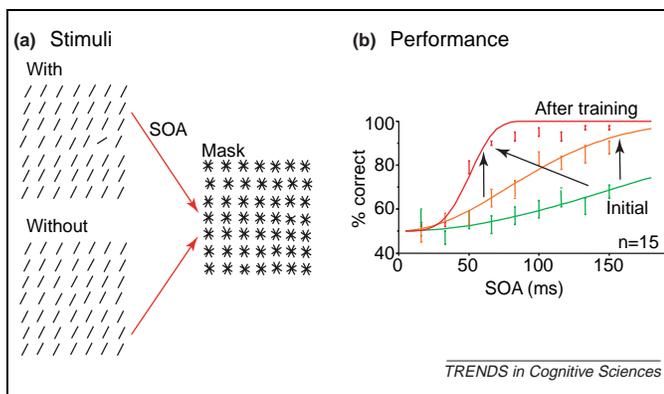

Figure 3. Learning dynamics for detection of an oddly oriented bar. Ahissar and Hochstein [17–19,21,22] focused on processes underlying improvement in detecting presence of an oddly oriented bar in an array of homogenously oriented distractor bars **(a)**. Detection would have been effortless, except that the stimuli were presented briefly (<250 ms) and quickly followed by a mask. Task difficulty was controlled by varying stimulus-to-mask onset asynchrony (SOA). **(b)** With practice, performance improved first for longer, then also for shorter SOAs. Early learning for long SOAs transferred to new orientations, suggesting that training first affects high-level generalized representations. When stimulus detection was made difficult, by using short SOAs, a small target/distractor orientation difference, or greater target position uncertainty and eccentricity, improvement was later and both orientation- and position-specific, suggesting that later learning reflects top-down guided low-level changes. This cascade of learning, from higher to lower areas, for easier to more difficult conditions, reflects the Reverse Hierarchy of perceptual learning.

in receptive fields along the visual hierarchy [21]. A similar easy–difficult specificity pattern was also found for inter-hemifield and cross-dimensional transfer [25,26], as well as for motion discrimination [27,28] (see also Box 2).

Following psycho-anatomy logic we concluded that easy task conditions (with large signal-to-noise ratios) are learned at higher cortical levels along the visual pathways, where receptive fields generalize across position and orientation. Difficult conditions, on the other hand, are learned at lower-levels where receptive fields are more specific to both retinal position and orientation. This scheme is illustrated in Figure 4.

Studying learning dynamics, we found that even though trials with different levels of difficulty (i.e. different SOAs) were presented in an interleaved manner, easy conditions (long SOAs) were learned first, and only after substantial improvement was achieved for these cases, did improvement begin for more difficult conditions (brief SOAs). Thus, in the temporal domain, learning proceeds in a cascade, from easier to more difficult conditions. From the psych-anatomy logic we deduced that learning begins at high-level areas and gradually progresses to lower-level areas when better signal to noise is needed. This cascade has ecological benefits, because high-level cortical representations are ecologically meaningful, and reflect the perceptual interpretation of the distal stimuli. However, their

Figure 4. Reverse Hierarchy Theory. Initial vision *at a glance* depends on high-level object and category representations built by implicit hierarchical processing [62]. In this way, initial high-level learning transfers over basic stimulus parameters [21]. Later vision with scrutiny is a return to simple feature details available at low levels [62]. Thus, later low-level learning is parameter-specific, being a low-level modification by guided return down the Reverse Hierarchy [21].
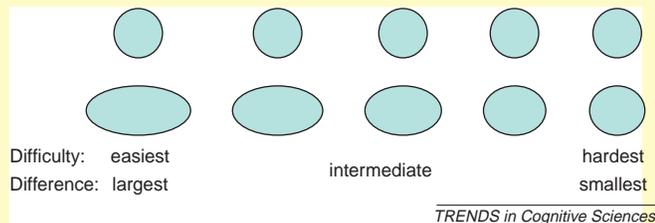
signal-to-noise ratio is poor, particularly with respect to spatial aspects, because generalized object selectivity is obtained in part by convergence across spatial parameters.

Learning is therefore attention driven, where attention is the mechanism for choosing the relevant neuronal population, by increasing its functional weight. It follows that initial high-level learning must precede low-level learning, as it provides the essential enabling stage for the backward search process.

The Eureka effect

An extreme condition occurs when subjects are trained only with difficult cases, so that they have no introductory easy trials along which to learn. Under these conditions, typically no learning occurs. However, a single exposure to the stimulus (as it appears on the screen) suffices to initiate an immediate learning process [21,29]. We termed the impact of this single exposure, which enables subsequent learning the Eureka effect (see Box 1). In terms of formal definitions, Eureka is a special case of priming. But whether such cases are indeed manifestations of the same underlying mechanisms is still unclear.

Theoretical and experimental challenges to RHT

The psycho-anatomy logic assumes a direct relationship between typical receptive field properties and the area's general function. This concept has been questioned [30] on grounds of cortical variability, claiming that although average receptive field size and orientation tuning broaden along the cortical hierarchy, substantial variability at any stage produces significant overlap between areas, including presence of small receptive fields at higher levels. Thus, even specific learning could stem from high-level modifications (discussed in [31]). This criticism in fact relates to the field's lack of knowledge regarding the exact relations between single receptive field properties and the area's computations (i.e. the neural code).
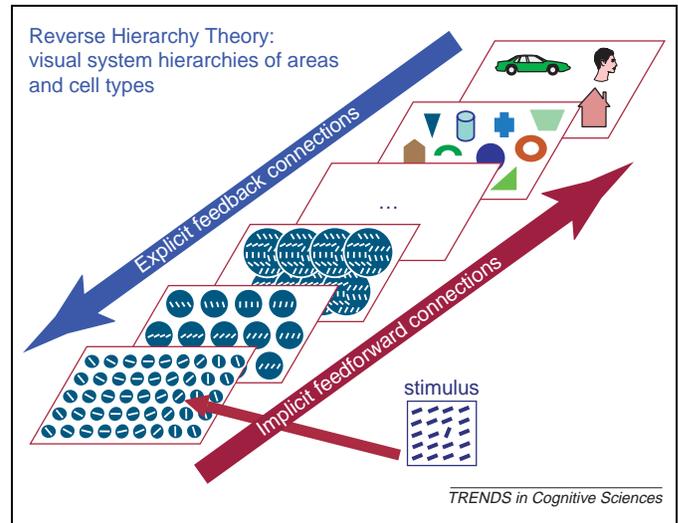
Currently, psycho-anatomical assumptions go beyond what has been directly substantiated. Still, the assumption that average receptive field properties denote the area's functional resolution is sensible, particularly when one assumes a population code. Moreover, the majority of findings in the visual learning literature are naturally accounted for by the RHT dynamic view.

An important RHT prediction is that perceptual learning will be contingent on task-specific attention, because the backward search for increased signal-to-noise ratio is attention-driven. Many studies, using 'pop-out' detection [17,32], orientation and texture discrimination [15,33], and Vernier acuity [34,35], are consistent with this prediction. However, recent studies in the visual [36,37] and tactile [38,39] modalities found that passive stimulation can also improve discrimination abilities.

In the visual modality, Watanabe *et al.* [36] found that when observers performed a difficult visual task near fixation while also exposed to surrounding moving dots, subsequent discrimination between the motion direction to which observers were continuously exposed and other motion directions was significantly improved. Thus, observers improved in motion discrimination even though the motion signal was not attended and had a sub-threshold coherence level (5%). Namely, it had not even been consciously perceived. Initial improvement was more specific to the local direction of motion and to its retinal position than subsequent improvement [40], suggesting a bottom-up dynamics of modifications. Both characteristics, namely bottom-up dynamics and no need for attentional control, have the flavor of adaptation processes (e.g. [41,42]). Such processes can lead to increased sensitivity to changes around the massively exposed stimuli (improved novelty detection). Still, the relations between this form of plasticity and practice-induced perceptual learning are not well understood.

It should be noted that RHT does not assert that there are no bottom-up induced modifications, such as adaptation

processes. It asserts that practising a perceptual task leads to a top-down cascade of weight retuning, which underlies the major bulk of improvement. Experimental evidence does suggest that when a task is performed, top-down processes dominate plasticity [43].

Electrophysiological support for a top-down cascade of learning

According to RHT, (successful) practice-induced plasticity at high areas should occur earlier and faster, and therefore be more prevalent and easier to detect than plasticity at low cortical areas. Indeed, evidence both from single neuron responses [44] and at larger scales, using fMRI [45], suggests that rapid, Eureka-like, long-term learning (Box 1) relates to modifications at higher-level visual areas (e.g. inferotemporal cortex, IT). Being a large-scale theory, however, RHT does not specifically predict the type of modification one should expect at the level of single neuron receptive field properties [46,47].

Several electrophysiological studies assessed training-induced modifications at different levels along the visual pathways and their basic findings are consistent with RHT predictions. First, IT representations rapidly tune towards task-relevant features and stimuli so that these become more salient [48–50] following behavioral time constants [51].

Second, according to RHT, task-related plasticity will gradually reach lower-level cortical areas. Thus, for overtrained monkeys, plasticity might reach V4 and subsequently lower-level areas, perhaps as low as V1 (although this would not be typically expected; [18–19]). Orientation discrimination is a task for which V1 is better suited than IT. Following initial training with a large orientation difference and modification of IT neurons, subsequent improvement should rely on lower-level neurons, with more plasticity in V4 than IT. This seems to be the case: following training with orientation discrimination, changes in IT receptive field properties were not orientation-specific – and were related to the behavioral context [52]. However, modifications in V4, leading to narrower tuning curves for orientation, were both orientation- and position-specific [30].

Evidence regarding training-induced plasticity in V1 is somewhat mixed, but the variability itself is consistent with RHT predictions that relate expected plasticity site to the specific behavioral paradigm applied. Thus, Schoups *et al.*, [53] trained monkeys on orientation discrimination to a level of expertise reaching just-noticeable differences (JNDs) of about 0.5 deg. Behaviorally, learning was position and orientation specific, accompanied by orientation and position specific plasticity at V1; (orientation tuning sharpened in neurons whose preferred orientation was ∼15 deg tilted from the trained one). Ghose *et al.* [54] also trained orientation discrimination, but found no evidence for V1 plasticity when applying the analysis of Schoups *et al.* [53]. The apparent discrepancy between these findings can be reconciled by RHT: Ghose *et al.* introduced variability along an irrelevant dimension (stimulus spatial frequency), inducing a large jitter in stimulus position. As low-level neurons are more sensitive to spatial parameters, unpredictable non-overlapping

low-level neuronal populations would be needed for each stimulus. Therefore, learning would not be expected to proceed 'so low'. Indeed, Ghose *et al.*'s monkeys never reached the expertise of the Schoups *et al.* monkeys; their asymptotic orientation JNDs were an order of magnitude larger, as was their learning generalization. All these are consistent with a higher-level plasticity site.

This comparison between monkey studies emphasizes RHT's prediction that the greater the stimulus variability during training (particularly when randomized), the lower the likelihood of low-level modification, achieving better thresholds, and stimulus specificity. Similarly, learning was generalized in humans practicing a difficult search task, with target and distractors randomly switched between trials, but attained thresholds were relatively high [55], reflecting high-level modification sites. Note that one can attain fine performance with reversed target and distractors, but practice has to be conducted in separate blocks [19].

The process of becoming an expert performer

Reverse Hierarchy Theory proposes the following complex pattern of learning as one extensively practices a given task and becomes an expert. Although the phases are described discretely, the transition is probably gradual:

(1) *Naïve performers* are governed by representations at the 'top' of the visual hierarchy, whether the task involves simple or complex perceptual attributes. When these representations are insufficient, mistaken or illusory, performance fails. 'Naïve performers' are actually common-sense experts, as their high-level representations are highly trained for ultra-rapid visual categorization of natural scenes. Indeed, Fabre-Thorpe *et al.* [56] found that such categorization is extremely rapid (e.g. detecting an animal in less than 150 ms) and cannot be further speeded by training.

(2) *Mildly-trained performers*, who have been exposed to a given task and context, show general improvement, consistent with high-level modifications and apparent in Eureka like phenomena, including establishment of top-down guidance and enabling mechanisms [57,21,29].

(3) *Highly-trained performers*, who have had a great deal of training experience, can manipulate the level they access for this tasks' performance. Under difficult conditions (i.e. when a better signal-to-noise ratio is needed) their performance is based on low-level representations and is thus specific to low-level aspects (e.g. [18–23,57,58]).

Their ability to access low-levels reflects changes of the integration properties of higher-level neurons, increasing the weights of task-relevant inputs and decreasing the weights (pruning) of inputs that are not informative for the task at hand. Thus, the effective tuning properties of these neurons are modified (V4 [30]; IT [50,59]) to improve signal-to-noise ratio (see [4] for model).

The changes in weights occur in a staged manner, beginning at the highest levels and gradually continuing at lower and lower levels along the reverse hierarchy. Top-down guidance is needed for these changes to occur. The particular neurons that modify are those whose outputs have previously been chosen as the relevant input to the next higher level. Thus, a chain of modified neurons is
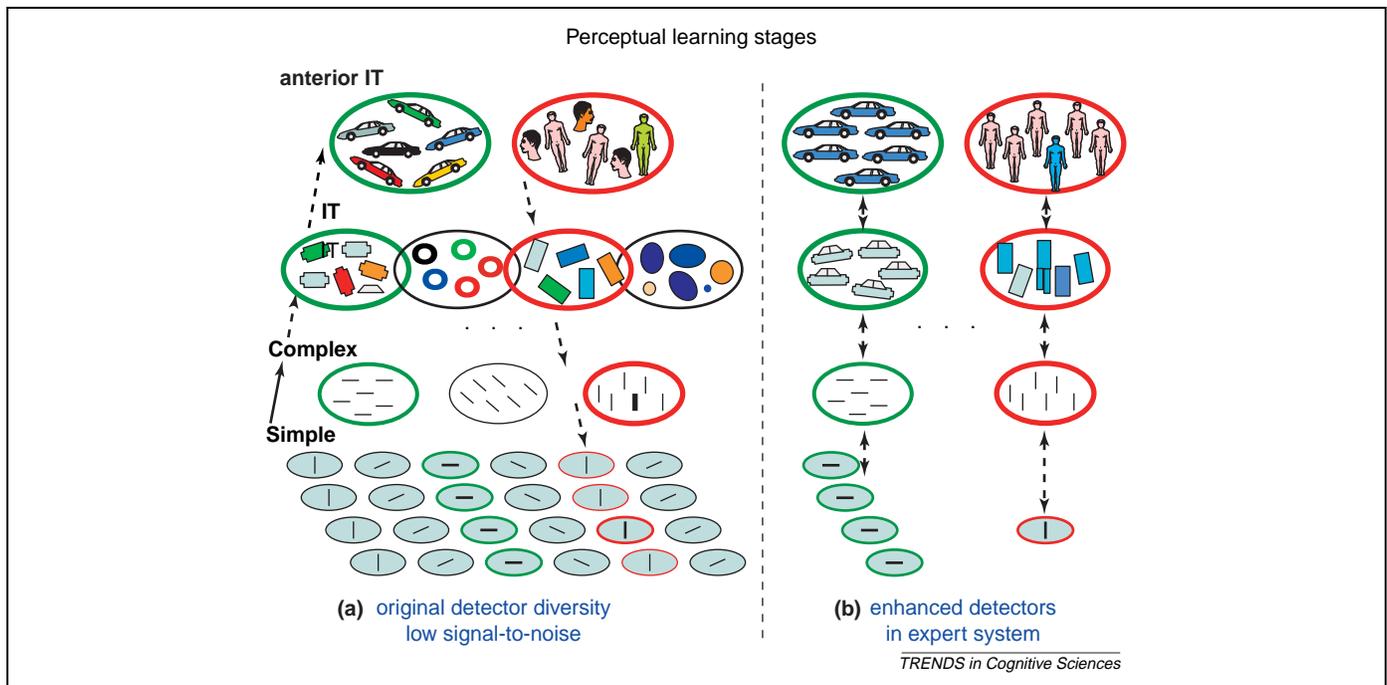
Figure 5. From naïve to trained to expert performance. **(a)** Naïve performance is based on activity at high-level cortical representations (which are formed by fast feedforward processing; arrows at left). This representation suffices for normal daily activity when scenes are simple and meet expectations. For difficult conditions and unexpected circumstances, conscious vision must seek details of the scene in lower areas, depending on top-down guidance over activated pathways. Training for these tasks enables subjects to perform this backward search more efficiently and to use the most appropriate level for the task on a trial-by-trial basis (dotted downward arrows). **(b)** Highly trained subjects have modified receptive fields first at higher and then also at lower cortical levels. Experts are those whose higher level representations have been modified by adding weight to appropriate inputs and pruning uninformative inputs (for the trained task). Performance can thus once again depend on higher, more generalizing, units.

progressively established which integrates task-relevant selected information at each level.

(4) *Expert performers*, who have had substantial amounts of training, can base their performance on higher levels again, even in difficult cases. As described above (highly-trained performers) there is a gradual process of modifications at lower and lower levels. Lower level modifications affect not only the levels of modification themselves, but also all higher levels that receive input from these levels. Consequently, higher-level representations are gradually better and better tuned for performance of the task as they attain an improved task-related signal-to-noise ratio. As there is always a preference for accessing the highest level that has sufficient task-related signal, higher-levels will be the basis for performance again. These levels have again the advantage of being more broadly tuned, because they are the outcome of convergence of multiple components. The increased weightings of informative inputs – at all cortical levels – now bias higher-levels towards the trained domain, which gains dominance compared with its emphasis in the original naïve distribution. A schematic illustration of this concept is shown in Figure 5.

This process yields the special characteristics of experts' perception. It is immediate and holistic on the one hand, suggesting high-level representations. Yet the generalization of their expertise to untrained conditions is limited, consistent with lower-level plasticity [60]. According to RHT, experts access high-levels again, but their high-levels are now biased towards the trained stimulus domain owing to the earlier top-down cascade of modifications.

The holistic nature of expert performance is captured in the classical term 'chunking' – referring to a process during which separate components gradually become a single perceptual or cognitive entity. A classic example is that of Morse decoders who gradually hear whole words rather than single bits. A timely example is that of expert video-game players who develop a spatially holistic mode of performance [61]. For example, they can enumerate a larger number of spatially distributed elements without counting. Interestingly, improved performance from video-game training occurred not only in the spatial domain but also in the temporal domain: video-game players had far less of an attentional blink than non video-game players, outperforming them even for lag one.

RHT attributes chunking and effortless performance to experts' ability to rely on high-level representations again. Even so, to explain the breadth of attentional changes characterizing expert performance [61], it should be further extended (see also Box 3).

RHT and perception

This article has focused on RHT and learning. However learning is not an odd case of perception. Rather it reflects the sequence of perception, attention and retention. Hence, RHT is expected to apply to perception in general. Specifically, learning begins at high-levels because these are the first levels accessed by conscious perception. Consequently, with brief exposures, we consciously perceive the 'gist of a scene', but not its fine details. Perceiving details requires access to lower levels, therefore more time and scrutiny [62]. We here summarize RHT's predictions for search tasks, which were the focus of our learning studies.

The main conceptual novelty of RHT in the context of visual search is the prediction that the determining factor of whether an odd element will effortlessly 'pop out' is its ecological relevance rather than its physical simplicity. This conclusion follows directly from the assumption that high-level representations are those that are immediately accessible to conscious perception. Because the function of high-level representations is to facilitate identification of ecologically relevant elements, those elements that are well segregated at these levels, will tend to pop out [24]. Although this seems at odds with the notion that only simple features pop out, it is highly consistent with the accumulated findings of complex features and high-level properties that in fact do pop out (reviewed in [24,62]).

So – why do simple features such as large orientation, size and color differences pop out? First, they only do so when they differ greatly from homogeneous distractors. Second, RHT's interpretation is that only those basic features that are relevant for object identification will pop out. Thus, a large orientation difference is important for telling apart vertically compared with horizontally elongated elements (e.g. horizontal cars versus vertical human figures, illustrated at the high-levels of Figure 2), red from green is important for segregating berries from their bush, and so on. Thus features will pop out if they form an ecologically relevant property, whether it is a simple property such as color or size, or a complex or conjunction property such as depth or shape or human face, classically termed 'emergent properties'. However, if the features or feature-combinations are not behaviorally relevant, they will have no (rapid) high-level segregated representation, and one would need to search at lower levels to find specific neuronal populations that respond to each aspect.

This interpretation predicts that searching for conjunctions that require serial scrutiny requires access to low-levels. This search would thus be hampered if activity at low-levels were disturbed during this later top-down scrutiny. By contrast, pop-out detection does not require re-access to lower-levels and would thus not be affected by a later disturbance to low-level activation. Indeed, such results were recently found [63,64] using repetitive transcranial magnetic stimulation (rTMS) that disrupted V1 activation 100–600 ms after stimulus presentation. Simple pop-out detection was not affected, whereas conjunction detection was significantly impaired. Recent electrophysiological results also support this conclusion, finding that behavioral pop-out is not correlated with single unit selectivity in V1. Although V1 neurons are selective to texture discontinuities, they are not more selective to orientation pop-out (single feature) than to a conjunction oddity (orientation and color) that does not pop out [65]. Taken together, both the micro and the macro scales are consistent with the concept that serial conjunction search and not parallel feature search, requires top-down access to low-level neurons, as predicted by RHT.

Conclusion

Reverse Hierarchy Theory proposes a 'neo-Gestalt' view of perceptual learning and of conscious perception in general, incorporating current knowledge of the anatomy, physiology and behavior of the visual system. RHT predicts that fine discriminations along simple dimensions best coded at low-level areas, will be difficult to learn because they require gradual back-tracking to neural populations we typically do not bother to access. Nevertheless, training with these tasks provides us with a window for studying basic mechanisms underlying the processes of visual learning and perception.

References
1 Gibson, E. (1969) *Principles of Perceptual Learning and Development*, Appleton Century Crofts
2 Fahle, M. and Poggio, T. eds (2002). *Perceptual Learning*, MIT Press
3 Ahissar, M. and Hochstein, S. (1998) Perceptual learning. In *Perceptual Constancies: Why Things Look As They Do* (Walsh and Kulikowski, J. eds), pp. 455–498, Cambridge University Press
4 Dosher, B.A. and Lu, Z.L. (1998) Perceptual learning reflects external noise filtering and internal noise reduction through channel reweighting. *Proc. Natl. Acad. Sci. U. S. A.* 95, 13988–13993
5 Barlow, H. (1990) Conditions for versatile learning, Helmholtz's unconscious inference and the task of perception. *Vision Res.* 30, 1561–1571
6 Dodwell, P.C. and Humphrey, G.K. (1990) A functional theory of the McCollough effect. *Psychol. Rev.* 97, 78–89
7 Ahissar, E. and Ahissar, M. (1994) Plasticity in auditory cortical circuitry. *Curr. Opin. Neurobiol.* 4, 580–587
8 Kapadia, M.K. *et al.* (1994) A quantitative measure for short-term cortical plasticity in human vision. *J. Neurosci.* 14, 451–457
9 Julesz, B. (1971) *Foundations of the Cyclopean Perception*, University of Chicago Press
10 Courtney, S.M. and Ungerleider, L.G. (1997) What fMRI has taught us about human vision. *Curr. Opin. Neurobiol.* 7, 554–561
11 Lerner, Y. *et al.* (2001) A hierarchical axis of object processing stages in the human visual cortex. *Cereb. Cortex* 11, 287–297
12 Avidan, G. *et al.* (2002) Contrast sensitivity in human visual areas and its relationship to object recognition. *J. Neurophysiol.* 87, 3102–3116

13 Hubel, D.H. and Wiesel, T.N. (1962) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.* 160, 106–154

14 Tanaka, K. (2003) Columns for complex visual object features in the inferotemporal cortex: clustering of cells with similar but slightly different stimulus selectivities. *Cereb. Cortex* 13, 90–99

15 Karni, A. and Sagi, D. (1991) Where practice makes perfect in texture discrimination: evidence for primary visual cortex plasticity. *Proc. Natl. Acad. Sci. U. S. A.* 88, 4966–4970

16 Poggio, T. *et al.* (1992) Fast perceptual learning in visual hyperacuity. *Science* 256, 1018–1021

17 Ahissar, M. and Hochstein, S. (1993) Attentional control of early perceptual learning. *Proc. Natl. Acad. Sci. U. S. A.* 90, 5718–5722

18 Ahissar, M. and Hochstein, S. (1996) Learning pop-out detection: specificities to stimulus characteristics. *Vision Res.* 36, 3487–3500

19 Ahissar, M. *et al.* (1998) Learning pop-out detection: building representations for conflicting target-distractor relationships. *Vision Res.* 38, 3095–3107

20 Schoups, A.A. *et al.* (1995) Human perceptual learning in identifying the oblique orientation: retinotopy, orientation specificity and monocularity. *J. Physiol.* 483, 797–810

21 Ahissar, M. and Hochstein, S. (1997) Task difficulty and the specificity of perceptual learning. *Nature* 387, 401–406

22 Ahissar, M. and Hochstein, S. (2000) The spread of attention and learning in feature search: effects of target distribution and task difficulty. *Vision Res.* 40, 1349–1364

23 Treisman, A.M. and Gelade, G. (1980) A feature-integration theory of attention. *Cogn. Psychol.* 12, 97–136

24 Wolfe, J.M. (2003) Moving towards solutions to some enduring controversies in visual search. *Trends Cogn. Sci.* 7, 70–77

25 Pavlovskaya, M. and Hochstein, S. (2003) Hemispheric specificity of perceptual learning effects under hard conditions. *J. Vis.* 3, 258a

26 Pavlovskaya, M. and Hochstein, S. (2004) Transfer of perceptual learning effects to untrained stimulus dimensions. *J. Vis.* 4, 263a

27 Liu, Z. (1999) Perceptual learning in motion discrimination that generalizes across motion directions. *Proc. Natl. Acad. Sci. U. S. A.* 96, 14085–14087

28 Lu, H. *et al.* (2004) Learning motion discrimination with suppressed MT. *Vision Res.* 44, 1817–1825

29 Rubin, N. *et al.* (1997) Abrupt learning and retinal size specificity in illusory-contour perception. *Curr. Biol.* 7, 461–467

30 Yang, T. and Maunsell, J.H. (2004) The effect of perceptual learning on neuronal responses in monkey visual area V4. *J. Neurosci.* 24, 1617–1626

31 Mollon, J.D. and Danilova, M.V. (1996) Three remarks on perceptual learning. *Spat. Vis.* 10, 51–58

32 Treisman, A. *et al.* (1992) Automaticity and preattentive processing. *Am. J. Psychol.* 105, 341–362

33 Shiu, L.P. and Pashler, H. (1992) Improvement in line orientation discrimination is retinally local but dependent on cognitive set. *Percept. Psychophys.* 52, 582–588

34 Levi, D.M. and Polat, U. (1996) Neural plasticity in adults with amblyopia. *Proc. Natl. Acad. Sci. U. S. A.* 93, 6830–6834

35 Herzog, M.H. and Fahle, M. (1997) The role of feedback in learning a vernier discrimination task. *Vision Res.* 37, 2133–2141

36 Watanabe, T. *et al.* (2001) Perceptual learning without perception. *Nature* 413, 844–848

37 Seitz, A.R. and Watanabe, T. (2003) Psychophysics: is subliminal learning really passive? *Nature* 422, 36

38 Godde, B. *et al.* (2000) Tactile coactivation-induced changes in spatial discrimination performance. *J. Neurosci.* 20, 1597–1604

39 Hodzic, A. *et al.* (2004) Improvement and decline in tactile discrimination behavior after cortical plasticity induced by passive tactile coactivation. *J. Neurosci.* 24, 442–446

40 Watanabe, T. *et al.* (2002) Greater plasticity in lower-level than higher-level visual motion processing in a passive perceptual learning task. *Nat. Neurosci.* 5, 1003–1009

41 Chaudhuri, A. (1990) Modulation of the motion aftereffect by selective attention. *Nature* 344, 60–62

42 Festman, Y. and Ahissar, M. Attentional states and the degree of visual adaptation to gratings. *Neural Netw.* (in press)

43 Recanzone, G.H. *et al.* (1992) Topographic reorganization of the hand representation in cortical area 3b owl monkeys trained in a frequency-discrimination task. *J. Neurophysiol.* 67, 1031–1056

44 Tovee, M.J. *et al.* (1996) Rapid visual learning in neurones of the primate temporal visual cortex. *Neuroreport* 7, 2757–2760

45 Dolan, R.J. *et al.* (1997) How the brain learns to see objects and faces in an impoverished context. *Nature* 389, 596–599

46 Crist, R.E. *et al.* (2001) Learning to see: experience and attention in primary visual cortex. *Nat. Neurosci.* 4, 519–525

47 Li, W. *et al.* (2004) Perceptual learning and top-down influences in primary visual cortex. *Nat. Neurosci.* 7, 651–657

48 Kobatake, E. *et al.* (1998) Effects of shape-discrimination training on the selectivity of inferotemporal cells in adult monkeys. *J. Neurophysiol.* 80, 324–330

49 Sakai, K. and Miyashita, Y. (1994) Neuronal tuning to learned complex forms in vision. *Neuroreport* 5, 829–832

50 Sigala, N. and Logothetis, N.K. (2002) Visual categorization shapes feature selectivity in the primate temporal cortex. *Nature* 415, 318–320

51 Jagadeesh, B. *et al.* (2001) Learning increases stimulus salience in anterior inferior temporal cortex of the macaque. *J. Neurophysiol.* 86, 290–303

52 Vogels, B. and Orban, G.A. (1994) Does practice in orientation discrimination lead to changes in the response properties of macaque inferior temporal neurons? *Eur. J. Neurosci.* 6, 1680–1690

53 Schoups, A. *et al.* (2001) Practicing orientation identification improves orientation coding in V1 neurons. *Nature* 412, 549–553

54 Ghose, G.M. *et al.* (2002) Physiological correlates of perceptual learning in monkey V1 and V2. *J. Neurophysiol.* 87, 1867–1888

55 Sireteanu, R. and Rettenbach, R. (2000) Perceptual learning in visual search generalizes over tasks, locations, and eyes. *Vision Res.* 40, 2925–2949

56 Fabre-Thorpe, M. *et al.* (2001) A limit to the speed of processing in ultra-rapid visual categorization of novel natural scenes. *J. Cogn. Neurosci.* 13, 171–180

57 Karni, A. and Sagi, D. (1993) The time course of learning a visual skill. *Nature* 365, 250–252

58 Fahle, M. (1994) Human pattern recognition: parallel processing and perceptual learning. *Perception* 23, 411–427

59 Logothetis, N.K. *et al.* (1995) Shape representation in the inferior temporal cortex of monkeys. *Curr. Biol.* 5, 552–563

60 Palmeri, T.J. *et al.* (2004) Computational approaces to the development of perceptual skills. *Trends Cogn. Sci.* 8, 378–386

61 Green, C.S. and Bavelier, D. (2003) Action video game modifies visual selective attention. *Nature* 423, 534–537

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

63 Juan, C.H. *et al.* (2004) Cortical interactions in vision and awareness: hierarchies in reverse. *Prog. Brain Res.* 144, 117–130

64 Juan, C.H. and Walsh, V. (2003) Feedback to V1: a reverse hierarchy in vision. *Exp. Brain Res.* 150, 259–263

65 Hegde, J. and Felleman, D.J. (2003) How selective are V1 cells for pop-out stimuli? *J. Neurosci.* 23, 9968–9980