Review
TRENDS in Cognitive Sciences Vol.8 No.8 August 2004
Full text provided by www.sciencedirect.com

Computational approaches to the development of perceptual expertise

Thomas J. Palmeri[1], Alan C-N. Wong and Isabel Gauthier[1,2]

Department of Psychology and Center for Integrative and Cognitive Neuroscience, 301 Wilson Hall, Vanderbilt University, Nashville, TN 37240, USA

Dog experts, ornithologists, radiologists and other specialists are noted for their remarkable abilities at categorizing, identifying and recognizing objects within their domain of expertise. A complete understanding of the development of perceptual expertise requires a combination of thorough empirical research and carefully articulated computational theories that formalize specific hypotheses about the acquisition of expertise. A comprehensive computational theory of the development of perceptual expertise remains elusive, but we can look to existing computational models from the object-recognition, perceptual-categorization, automaticity and related literatures for possible starting points. Arguably, hypotheses about the development of perceptual expertise should first be explored within the context of existing computational models of visual object understanding before considering the creation of highly modularized adaptations for particular domains of perceptual expertise.

Expertise has various manifestations. Experts differ from novices in areas such as solving physics problems, playing chess, making medical diagnoses, and performing athletic movements. The development of expertise in real-world domains involves a complex interplay of changes in perception, categorization, memory, problem solving, coordination, skilled action, and other components of human cognition. Here, we focus on some important aspects of *perceptual* expertise, particularly the changes that occur in how people perceptually categorize, identify and recognize visually similar objects.

Early expertise research focused on problem solving, decision making and reasoning. But recent years have seen growing interest in perceptual expertise [1–4]. Understanding perceptual expertise is more than characterizing the behavior of individuals with idiosyncratic skills in highly specialized domains. Perceptual expertise is an alternative to modular accounts of face recognition [1,5,6] and word recognition [7]; certain classes of objects might be 'special' not because of their intrinsic status, but because we have expertise with them. Perceptual expertise might tune perceptual representations; perception might not just passively describe the world, but change according to experience and goals [8–10]. Perceptual expertise could be at the core of seemingly abstract knowledge in an expert domain; conceptual understanding might be grounded in specific perceptual experiences rather than abstract amodal representations [11]. Perceptual expertise might be the endpoint of the normal learning trajectory; progress in understanding cognition can be accelerated by examining the extremes, so understanding perceptual expertise should help us understand how we interact with objects generally.

Formal computational models can characterize mechanisms underlying how we represent perceptual information, how we represent object knowledge, and how we make decisions about objects, such as recognizing, identifying and categorizing them. Computational models build on verbal descriptive theories by specifying representations and processes using formal mathematics and computer simulations. This forces the theorist to be explicit about all aspects of information processing, allowing theories to be clearly articulated, rigorously evaluated and potentially falsified [12]. Both the 'object recognition' and 'perceptual categorization' literatures have seen the development of sophisticated models that can account for detailed aspects of visual object understanding [8]. Many models account for learning quite generally, and some have specifically addressed aspects of perceptual expertise. This article does not propose a comprehensive computational model of the development of perceptual expertise, but instead looks to existing models for potential hypotheses. It should be emphasized that each of the computational approaches discussed in this review has been precisely formalized in mathematical equations or computer simulations. Because of the space limitations of this review, we cannot describe the details of these computational formalisms but urge the reader to examine the cited articles for detailed descriptions.

The development of perceptual expertise

It goes without saying that experts know more than novices. They can verbalize more properties, describe more relationships, make more inferences, and so forth [13–15]. That is what makes them experts after all. Our focus in this review is on how expertise is manifest in more perceptually oriented tasks.

Review of behavioral characteristics

Any comprehensive theory of the development of perceptual expertise must account for several behavioral characteristics that distinguish novices and experts:

(i) Novices often rely on verbal knowledge in the form of rules or ideal cases acquired from reference manuals or

Corresponding author: Thomas J. Palmeri (thomas.j.palmeri@vanderbilt.edu).
[1] Learning Sciences Institute.
[2] Vanderbilt Vision Research Center.
Available online 2 July 2004

 1364-6613/$ - see front matter doi:10.1016/j.tics.2004.06.001

explicit instruction, or created via induction. By contrast, although experts have more verbal knowledge about a domain, expert categorization often seems removed from explicit deliberation [15].

(ii) Novices are slow and deliberate, perhaps reflecting their use of explicit rules. The development of expertise is accompanied by marked increases in speed of performance, originally characterized as the power law of practice [16], but more recently characterized by other mathematical functions [17,18].

(iii) Novices are easily distracted, whereas experts can engage in other tasks while making expert judgments. But perhaps that apparent lack of interference results from experts no longer using verbalizable routines – concurrent verbal activity does not interfere because expertise is not dependent on verbally mediated processes. In fact, expert judgments made while observers engage in concurrent tasks that tap the same representational resources suffer interference that is not found for novice judgments [1,19,20].

(iv) For novices, judgments at a basic level ('dog') are faster than judgments at either a superordinate ('animal') or subordinate level ('terrier'). The fastest level of categorization is often described as the 'entry-level' into conceptual knowledge [14,21]. For experts, there is an 'entry-level shift' whereby subordinate categorizations are as fast as basic-level categorizations [14,22].

(v) Novices can often attend to part of an object and ignore irrelevant parts. By contrast, experts often show less flexibility in selectively attending to relevant parts and suffer interference from irrelevant variation. For example, in a part-matching task – adapted from work in face recognition [23] – subjects are asked to attend to the top part of a whole object. After a delay, a second object is shown with the irrelevant bottom part either matching or mismatching the bottom of the first object. When judging whether the top is the same or different, novices are unaffected by the irrelevant bottom, whereas experts show facilitation when the bottom matches the first object and interference when it mismatches [24].

(vi) In other situations, novices are unable to parse an object into the appropriate parts. By contrast, experts have learned to dimensionalize a complex object in a way that is consistent with the goals of their expertise [25,26].

(vii) Experts are sensitive to changes in the configuration of features (e.g. switching the relative positions of the mouth and the eyes), but only when objects are presented in a familiar orientation [19,27]. By contrast, the ability to detect featural changes (e.g. a change in the shape of the mouth) is less dependent on orientation, and some evidence suggests that it develops earlier [28].

(viii) Experts generalize their knowledge. They can learn to categorize new objects more quickly than novices, so long as the new objects are like other objects in their domain of expertise, systematically varying in the same qualitative manner [19].

(ix) However, the generalizing ability of experts is limited. For example, experience is often limited to particular viewpoints. In the same way that face recognition is impaired by inversion, expert object recognition is also impaired by inversion [5,29]. Indeed, the development of automaticity often shows limited generalization [30].

No single computational theory from the object recognition or perceptual categorization literatures can account for all aspects of the development of perceptual expertise (see [8]). However, some existing models speak to certain aspects of expertise, and a comprehensive theory might be possible by combining complementary models. This article sketches some possible avenues for such integration.

A 'simple' model

We first consider a 'simple' model of the development of perceptual expertise. Recall that objects are categorized fastest at an intermediate entry level for novices [14,21]. Perhaps fastest means first: an initial stage categorizes an object at its entry level whereas later stages categorize at more subordinate or superordinate levels. With expertise, a new entry level is established, so that objects are categorized as quickly at a more subordinate level [14]. But what does it mean to 'establish a new entry level'? It could mean creating a new special-purpose module (or perceptual routine [21]) for categorizing objects at a subordinate level. The creation of such a module would account for the automaticity, domain-specificity and attentional inflexibility (cognitive impenetrability) seen with perceptual expertise.

Although intuitive, this 'simple' model is limited in important ways. First, simply positing a new module for entry-level categorization begs the question of how and why a new module might be created, and how this module works. More generally, such modular solutions to object recognition have been challenged by neuropsychological [31], neurophysiological [32] and neuroimaging [2,33,34] results. The need for separate modules for different levels of categorization has also been challenged on computational grounds [35,36,37]. Arguably, it would seem important first to try to ground explanations of perceptual expertise in existing computational models that account for how novices recognize, categorize and identify objects. If such explanations are found wanting, then highly specialized, modular adaptations need to be investigated.

Computational models of the development of perceptual expertise

We focus on models from the object recognition and perceptual categorization literatures, two fields of visual object understanding that grew from largely separate research traditions but have recently begun to converge empirically and theoretically [8]. As this is a selective review, we necessarily omit several important alternative theoretical approaches [38,39], but our selection was aimed at presenting a coherent theoretical package. Figure 1 outlines the relationships between the various theoretical approaches and specific computational models that are cited in this review.

Aspects of several successful computational models are illustrated in Figure 2. A hierarchy of early and mid-level visual processing stages transforms the high-dimensional retinal image into a relatively low-dimensional object representation in terms of shape, color, size and other perceptual dimensions [40] ('early form processing' and 'other visual processing' in Figure 2). Illustrated for early form processing is a recent proposal for creating

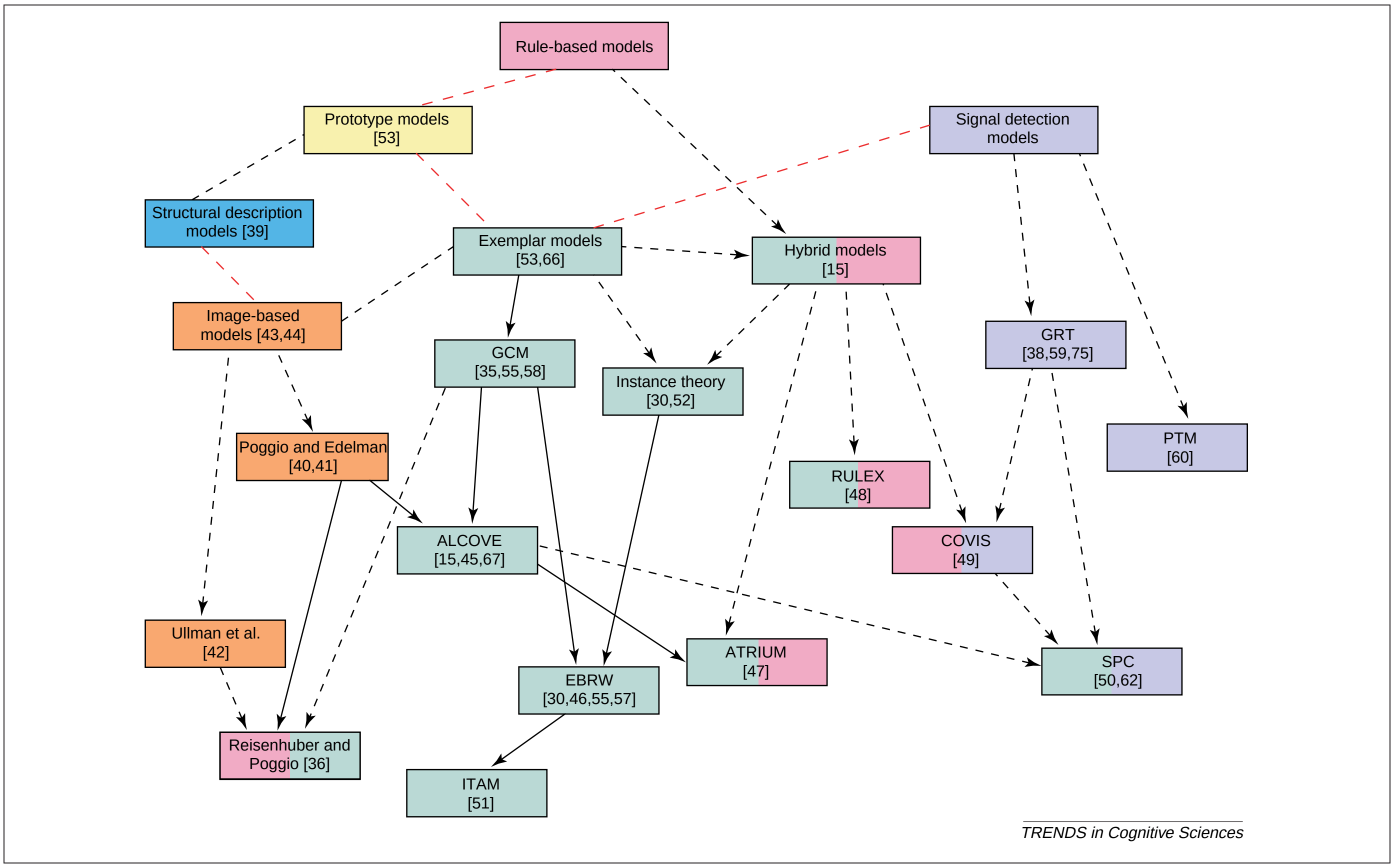

Figure 1. **Relationships between cited computational models**. Relationships between the various computational models cited in this selective review. General classes of models (color-coded) include structural description (blue) and image-based (orange) models from the object-recognition literature, and rule-based (pink), prototype (yellow), exemplar (green), and signal-detection (purple) models from the perceptual categorization literature (see [8]). Specific models are color-coded according to their basic class; models combining aspects from two classes are color-coded accordingly. The figure is ordered roughly chronologically from top to bottom. Models within a direct lineage are indicated by solid lines. Models having a more indirect influence are indicated by dotted lines. Classes of models that have historically been contrasted are connected by red dashed lines. GCM = Generalized Context Model [35], ALCOVE = Attention Learning COVEring map [45], EBRW = Exemplar-Based Random Walk [46], ITAM = Instance Theory of Attention and Memory [51], RULEX = RULE-plus-EXception [48], GRT = General Recognition Theory [38], COVIS = COmpetition between Verbal and Implicit Systems [49], PTM = Perceptual Template Model [60], SPC = Striatal Pattern Classifier [50].

translation- and size-invariant representations through a hierarchy of filters of increasing size and complexity [36]. According to image-based theories of object recognition [40], object shape has a distributed representation through activation of units selectively tuned to particular views of a whole object ('view-tuned units') [36,40,41] or stored views of components (image-based parts) of an object ('component-tuned units') [36,42]; alternative theories propose that objects are represented in terms of three-dimensional primitives such as geons [39]. Image-based theories provide accounts of the viewpoint-dependent nature of visual object recognition [38,43,44], and how apparent viewpoint-independent recognition might be possible when a large collection of views is stored.

To *recognize* an object is to decide that its perceptual representation is similar to a representation created during a previous experience and stored in memory. But to *identify* or *categorize* an object, its perceptual representation must be compared with a knowledge representation that summarizes what is known about the identity or category of objects. This knowledge mediates the association between a perceptual representation and some form of internal or external act (carried out by 'decision units' in Figure 2), the simplest of which might be naming the object [36,45,46].

One possibility is that category knowledge is represented in terms of explicit, perhaps verbal, rules ('rule system' in Figure 2) [15,47–49]. A rule typically involves independent assessment of individual dimensions, with the simplest rule being something like '*big things are members of category A*' (Figure 3a, left panel). Another possibility is that a representational system, variously described as implicit and procedural [50], object-based [36], or exemplar-based [35,46,45,51], is used to categorize (or identify) an object. 'Exemplar units' are tuned to particular combinations of form, color, size and so forth; in effect, the receptive fields of exemplar units span a small region of the representational space created by mid-level vision [38,45] (Figure 3, right panel). In the extreme, each stored exemplar is associated with a particular experience [35,52]; alternatively, exemplar units could be more coarsely distributed in a manner that is influenced by, but not related in a one-to-one way to, experience [45,50,53]. An important component of some exemplar models is that selective attention ('α nodes' in Figure 2) highlights diagnostic dimensions [35,45,54], with diagnosticity varying according to the demands of the task [55]. Exemplars are mapped with learned strength to units associated with particular categorizations or identifications in a decision layer, generically referred to as

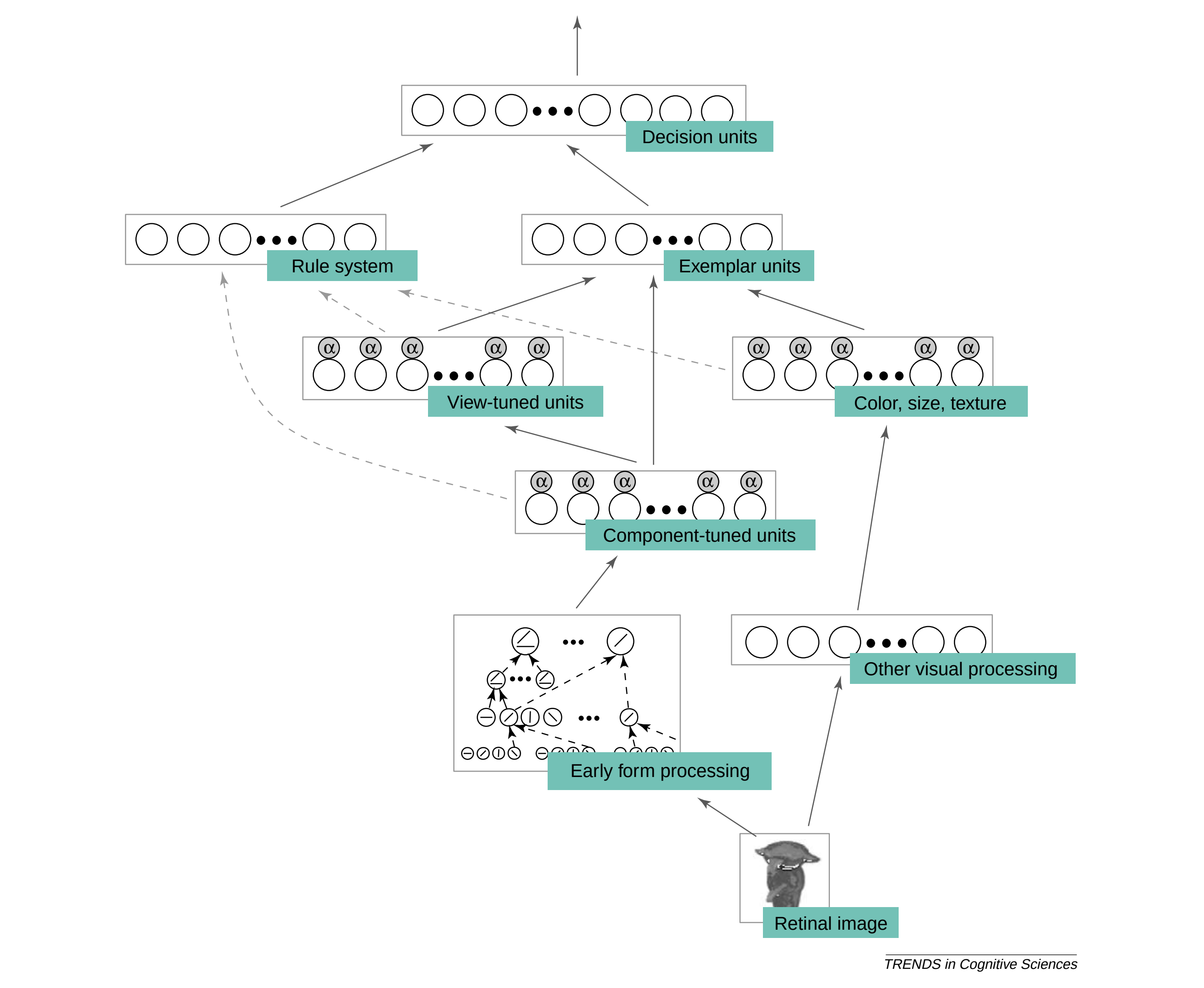

Figure 2. Summary of stages of visual information processing. The processing stages across several models from the object-recognition and perceptual-categorization literatures (see [8]). Starting with a retinal image, a hierarchy of processing stages ('early form processing' and 'other visual processing') transforms the retinal image into a relatively low-dimensional object representation in terms of form, color, size and so forth. According to a class of image-based theories from object recognition, an object's shape is represented in terms of image-based components (component-tuned units) or entire views (view-tuned units). According to several current theories, the perceptual representation of an object can be associated with a category (or its identity) through (at least) two forms of knowledge representation: a rule-based (rule system) or exemplar-based or exemplar-like system (exemplar units). A decision stage (decision units) weighs the evidence for particular categorizations (or identifications) and chooses a response. An important component of many models is that selective attention (α nodes in the figure) can highlight aspects of the perceptual representation that are particularly diagnostic for a decision.

decision units. This decision stage selects a response based on evidence for various categorizations or identifications, with various proposals offered for how this decision process operates [46,56,57]. The decision stage must also resolve any competition between the rule system and the exemplar system, again with various proposals offered for how this competition is resolved [18,30,47,49]. To be clear, the illustration in Figure 2 shows only feedforward connections. Top-down feedback connections also exist to implement task-dependent dimensional selective attention and to guide learning at various stages of object processing.

We will now articulate some learning-related changes in the visual-object understanding system that could accompany the development of perceptual expertise. In organizing these hypotheses, we begin by describing potential changes that could occur relatively late in the system (i.e. the top of Figure 2) and work our way back to changes that could occur at earlier levels (bottom of Figure 2).

Rules-to-exemplars shifts

Theories of the development of automaticity can provide insight into the development of perceptual expertise. Logan linked automatic processes with memory processes [51,52]. According to his 'instance theory', tasks can be performed either by using explicit rules or by retrieving a solution from memory, whether that task is solving a physics problem or categorizing an object. Formally, the model assumes a horse race between rule use and memory retrieval. Because novices have insufficient experience to build up representations for past solutions, they rely on explicit rules. Exemplar memories associated with experiences are gradually stored, and then as memory is

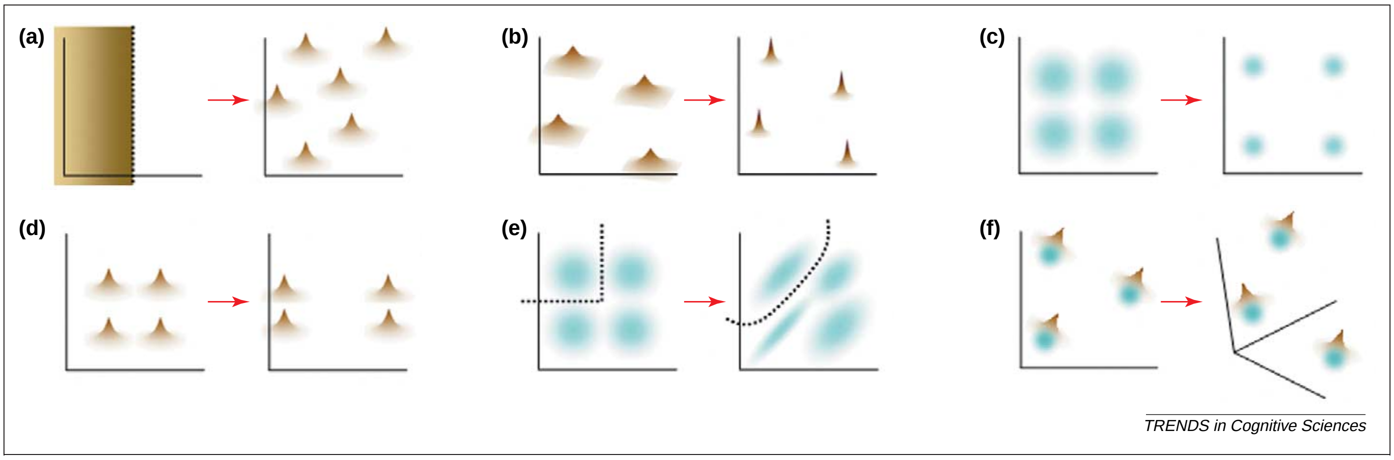

Figure 3. **Some possible representational changes**. Some of the various possible representational changes that can occur with the development of perceptual expertise. For each pair of panels, the left one indicates early learning (novice) and the right one indicates later learning (experts). Knowledge representations are shaded brown; perceptual representations are shaded blue-green. **(a)** A simple unidimensional rule is used to represent a category early in learning, but localized exemplars are used to represent a category later in learning. **(b)** Exemplars have broadly tuned receptive fields early in learning, but have more finely tuned receptive fields later in learning. **(c)** Perceptual noise is quite large early in learning, but perceptual noise is smaller and representations are more discriminable later in learning. **(d)** Dimensions are equally attended early in learning, but diagnostic dimensions are selectively attended and nondiagnostic dimensions are unattended later in learning. **(e)** During learning, perceptual representations can come to violate perceptual independence (as indicated by correlated noise between dimensions), perceptual separability (as indicated by the dependency relations of perceptual noise between dimensions), or decisional separability (by placing a decision boundary non-orthogonally to a dimension). **(f)** Perceptual representations can change qualitatively by combining existing dimensions or creating new dimensions.

strengthened, retrieval eventually becomes faster than rule use. So, according to instance theory, the development of automaticity and expertise reflects a shift from rule use to memory retrieval. Automatic processes are fast and obligatory because memory retrieval is fast and obligatory. Automatic processes no longer follow typical metrics of problem difficulty because memory simply encodes an object with its response. And finally, automatic processes show limited generalization to the extent that memories for particular instances show limited generalization.

Recently, generalized and extended versions of the instance theory have linked categorization with automaticity [30,46], and other theories have proposed a competitive choice between an algorithm and memory retrieval rather than a race [18], or combined contributions from both [47,49]. In these computational models, the development of expertise involves a shift from rule-based to exemplar-based representations. The automatic nature of perceptual expertise, including the non-verbal nature, the faster response times, and limited generalization to new objects, naturally fall out of these various models [30]. But these models, at least as currently formulated, are probably insufficient to account for many of the perceptual effects seen with expertise.

Increases in memory sensitivity

According to Logan's instance theory, memory retrieval gets faster because more and more exemplars have been stored. But in addition to increasing the strength of memory representations, there can also be an increase in the quality of those representations – what might be termed memory 'sensitivity'. Early in learning, retrieval can be hampered by memory noise [58] interfering with successful retrieval. Later in learning, memory representations can become more finely tuned and less susceptible to noise. Nosofsky proposed increases in memory sensitivity to account for learning in experiments where subjects learned to uniquely identify similar objects [59].

The activation of an exemplar is a function of its similarity to the perceptual representation of an object. If you think of an exemplar representation as having a receptive field that covers a local region of a multi-dimensional perceptual space, then it is activated to the extent a perceptual representation fall within its receptive field; emerging neural evidence supports important aspects of such hypothesized representations (see [8,38]). Early in learning, exemplars are broadly tuned, such that a range of perceptual representations activate the same exemplars, but later in learning, receptive fields become finely tuned (Figure 3b). In addition, the location of the exemplars can change over learning to better reflect the true or optimal location of exemplars [60]. Perceptual expertise might be associated with memory representations having high fidelity and that veridically reflect experienced objects [35] or that optimally map objects onto learned categories or identities [50].

Increases in perceptual sensitivity

An exemplar representation mediates the mapping of a perceptual representation of an object to a recognition, identification, or categorization of that object. The quality of that mapping is partially determined by the quality of the exemplar representation, but also partially the quality of the perceptual representation. Perception is inherently noisy [38,61]: the same object can give rise to a different perceptual representation from trial to trial. Arguably, perceptual noise has relatively little influence when stimuli are clearly discriminable from one another [62]. But perceptual noise can have a marked influence on performance when fine distinctions are required [63], such as when very similar objects must be uniquely identified.

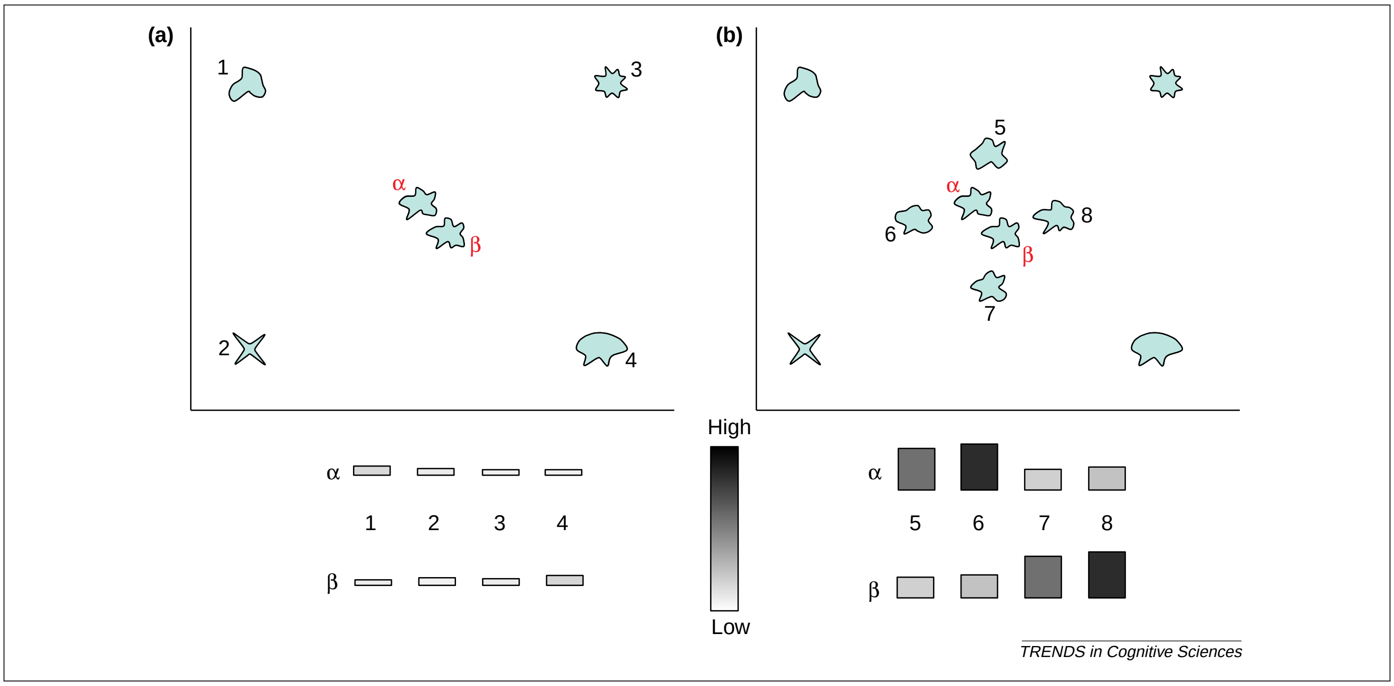

Figure 4. Representation of novel shape by image-based models. According to image-based models, objects are represented by their similarity to stored views of other objects [36,41,40]. An experienced view of an experienced object might have an explicit memory representation, but novel views of experienced objects or views of novel objects have an implicit or 'ephemeral' representation in terms of similarities to previously experienced objects [40]. Image-based models provide a possible way of thinking about increases in perceptual sensitivity with the development of perceptual expertise. **(a)** Let us assume that objects 1–4 have been previously experienced and have explicit image-based representations. Novel objects α and β are represented according to their similarities to stored images. Because α and β are relatively similar to one another but relatively dissimilar to any stored image, their distributed representation of similarities to stored images is extremely similar, making discrimination difficult. The bottom part of the panel depicts representations for novel objects α and β in terms of their similarity to stored images for objects 1–4 (with similarity given by both the height and the shading of the bars). **(b)** Now objects 5–8 have been explicitly represented in memory after experience with a homogenous class of objects. Novel object α and β are again represented according to their similarities to stored images. But because there are now stored images very similar to the novel objects, the distributed representations of those novel objects (below) are quite dissimilar, making discrimination easy. Perceptual sensitivity can be increased when additional images of objects populate a region of shape space.

Perceptual expertise could cause a decrease in perceptual noise for objects in a domain (Figure 3c).

Perceptual learning during the development of expertise can lead to perceptual representations that are more distinct from one another [9]. Perceptual distinctiveness might be more marked at boundaries separating objects from different categories. This predicts an acquired categorical perception for objects [64] analogous to that seen for phonemes in speech [65]. But perceptual distinctiveness might also extend to other areas of perceptual space beyond a category boundary [66]. Indeed, expertise generalizes to new objects within the space of previously learned objects [20] and this might require a local retuning of perceptual representations.

A variety of simple perceptual learning mechanisms can lead to enhancement of perceptual representations along basic dimensions such as orientation, hue, brightness and so forth. But perceptual learning of more complex distinctions between objects might require something different. In particular, complex novel objects pose an interesting problem to observers, because they can be dimensionalized in many possible ways. Image-based models provide one possible mechanism for how enhanced perceptual representations of novel stimuli can emerge with the development of perceptual expertise. As illustrated in Figure 4, novel stimuli have a distributed representation in terms of similarity to stored views in existing image-based models [40,36]; as the density of stored views within a space increases, greater representational disciminability might be possible.

Changes in selective attention

According to some models, an important aspect of learning to identify and categorize objects is learning to attend to diagnostic dimensions that maximally distinguish objects [35,45,57] (see Figure 3d). For relatively simple objects with a clear dimensional structure, this selective attention can occur relatively quickly, perhaps within just a few dozen training trials [67]. But for objects with multiple dimensional interpretations, learning dimensional selective attention might be more protracted [68]. Moreover, with minimal training, selective attention can be quite malleable, but with extended training, patterns of selective attention can become manifest in the very perceptual representations that support categorization, perhaps to the level of neurons representing objects in inferotemporal cortex [54,69].

Perceptual learning can be cast in terms of learned selective attention. Expertise could involve more densely populating a dimension with additional perceptual detectors, or existing perceptual detectors could become more finely tuned. But a simpler mechanism does not require creating new representations or retuning existing representations: perceptual learning could instead involve learning to attend selectively to perceptual detectors tuned to the region of a dimension that is most informative for making a discrimination [60]. Finer discriminations are possible by ignoring irrelevant detectors in favor of those that are maximally informative.

To the extent that patterns of selective attention become automatically established, or that patterns of selective attention become hardwired in object representations [54], perceptual expertise can introduce inflexibility relative to novice performance. To be sure, learned selective attention benefits experts for identifying or categorizing objects, but expertise can introduce a cost when tasks violate this learned selective attention, such as the need to ignore a part of an object that would otherwise be attended to.

We should emphasize that changes in memory sensitivity, perceptual sensitivity and selective attention are not passive processes. Behavioral evidence suggests that perceptual categorization, not perceptual exposure *per se*, is important for the development and generalization of perceptual expertise [3]. Indeed, most existing models assume that learning is strongly driven by top-down factors in response to corrective feedback [9,15,18,45,49,67,25].

Changes in perceptual representations

Finally, qualitative changes in the underlying perceptual representations are perhaps the most complex kinds of changes that could emerge with the development of perceptual expertise (Figure 3f).

In one characterization, becoming an expert means acquiring the right repertoire of parts, features or dimensions for representing objects in a way that serves the goals of that expertise. In some cases, complex objects have continuous dimensions that are hard to describe or have multiple possible (competing) dimensionalizations. Category learning can have a powerful influence on the way such complex objects are dimensionalized ([25,68], but see also [70]); once appropriately represented, the most diagnostic dimensions can be selectively attended to maximize categorization performance. In addition, complex objects can also be parsed onto many possible (competing) parts or features. Category learning can create a new vocabulary of perceptual components for describing complex objects in the service of challenging categorizations [10]. Progress is being made developing models that create novel features in response to category feedback [25].

In another characterization, novices represent objects in a part-based manner and experts represent objects holistically [24]. According to some image-based theories, objects are represented at an intermediate stage in terms of view-based parts of intermediate complexity [36,42] (Figure 2). These are not parts (such as geons) as described by structural description models [39] but are two-dimensional images of components of a complex object. Novices could represent novel objects by a distributed representation of such image-based components, and these representations might be sufficient to support categorization at an intermediate level [42]. However, such representations might be insufficient to support more fine-grained categorization or identification. Novel objects could be represented according to their similarity to image-based representations of full objects [40], although in some cases stored images of experienced objects might just be too dissimilar to allow distributed representations of novel objects that are sufficiently discriminable from one another (see Figure 4). Thus, the demands of fine-grained discrimination required for expertise could cause the creation of image-based representations for novel objects. This might not occur in one fell swoop, but there could be a gradual creation of larger and larger image-based components [1,71] and the creation of such components could be guided by top-down influences [10]. Therefore, a shift from part-based to image-based representations need not mean a shift from a part-based to an image-based representational system. Rather, all objects could be represented within a hierarchy including image-based components at early stages and full images at later stages, with the demands of the task dictating which level is most relevant, and with learning influencing when and if new full image-representations are created.

Although various holistic effects have been repeatedly documented [1,19,72], unequivocal evidence for the creation of holistic representations has been rather mixed. There is evidence for the development of unitized perceptual representations [73,74]. However, evidence for a shift from dimensions processed in a separable manner to dimensions processed in an integral manner has been challenging to document [73]. As illustrated in Figure 3e, holistic representations have sometimes been operationalized in terms of violations of ‘perceptual independence’ between dimensions (i.e. that perceptual noise is correlated between dimensions) and violations of ‘perceptual separability’ between dimensions (i.e. that the perceptual noise along one dimension depends on the value of another dimension) in the context of multidimensional signal detection theory [38]. Some recent work suggests that neither violation is found with faces, the most salient example of stimuli that are represented holistically; holistic effects might instead emerge because dimensions are integrated at a decisional stage of processing, a violation of ‘decisional separability’, rather than being perceptually integrated in their representations [75]. Thus, although learning mechanisms might exist for the development of holistic representations with the development of perceptual expertise, evidence for the creation of such holistic representations remains a source of some debate and much current research.

Concluding remarks

Expertise could entail the creation of highly specialized, modular adaptations to an object domain. However, we argue here that it is important first to try to ground explanations within existing computational models that already account for important aspects of visual object understanding. After reviewing current models from the object-recognition and perceptual-categorization literatures, we identified several hypotheses for the development of perceptual expertise. Some of these hypothesized changes occur at high levels of the system (shifts from rule-based to exemplar-based representations), whereas others occur at low levels of the system (fundamental changes in perceptual processing). Although learning-related changes can take place throughout the cortical pathway,

Box 1. Questions for future research

- Can aspects of the computational models outlined in this review be combined into a comprehensive model of perceptual expertise? We have outlined how existing computational models from the object-recognition and perceptual-categorization literatures account for important aspects of the development of perceptual expertise, but no single model can account for all key phenomena.
- Might image-based models from object recognition and exemplar-based models from perceptual categorization provide a natural theoretical integration? These theories adopt complementary assumptions about the nature of the representations, organization and goals of visual object understanding [8,36,37].
- Are there neural correlates of the representations and processes instantiated in the various computational models? Some emerging results are quite compatible with many of the computational models we outlined here [8,32,36,37,54,76,77].
- How can relatively abstract computational models be related to the realities of cortical organization; how is cortex organized to represent objects and object categories, and can current models be extended to manifest the anatomical and physiological constraints on information processing?
- Can existing computational approaches account for how different kinds of expertise might emerge [78]? For example, contrast the expertise uniquely identifying similar faces from the expertise identifying letters irrespective of quite dissimilar fonts [79].
- What is the time course of categorization, identification and recognition decisions throughout the development of perceptual expertise? Extant models from the object-recognition and perceptual-categorization literatures have largely focused on response probabilities, but some models account for response times as well [30,46,51,66,77]. Can these models account for emerging details of how object decisions evolve over time [22,66,80]?
- If there are multiple systems for categorizing, identifying and recognizing objects – whether those systems be modular or highly interactive – what is the most appropriate and most parsimonious characterization of the nature of those systems and their interactions [11,15,18,34,49,76,79]?

changes might occur more easily, and on a faster time scale, at later levels of the system – creating memories, selective attention, and new decision criteria. The earliest stages of visual object processing exhibit marked plasticity both evolutionarily and developmentally, but there might be resistance to drastic changes early on in the perceptual system, because of the expected stability of the world we experience. Therefore, from a theoretical standpoint, it seems fruitful to examine the sufficiency of simpler changes in memory, attention and decision-making before considering whether early visual-perceptual processing need be fundamentally modified. That said, if the same behavioral phenomena of expertise can be explained by changes at multiple levels of the system, then neurophysiological evidence will need to be brought to bear to resolve this challenge (see also Box 1).

Acknowledgements

The authors' work is supported by NSF Grants BCS-0218507, BCS-9910756 and BCS-0091752, NIMH Grant R01 MH61370, NEI Grant R01 EY13441, NEI Grant P30 EY008126, and a grant from the James S. McDonnell Foundation. The authors wish to thank members of the Perceptual Expertise Network (funded by JSMF) for helpful discussions.

References

1 Gauthier, I. and Tarr, M.J. (2002) Unraveling mechanisms for expert object recognition: bridging Brain activity and behavior. *J. Exp. Psychol. Hum. Percept. Perform.* 28, 431–446
2 Tanaka, J.W. and Curran, T. (2001) A neural basis for expert object recognition. *Psychol. Sci.* 12, 43–47
3 Tanaka, J.W. *et al.* The training and transfer of real-world, perceptual expertise. *Psychol. Sci.* (in press)
4 Tanaka, J.W. (in press) Object categorization, expertise and neural plasticity. In *The New Cognitive Neurosciences* (3rd edn) (Gazzaniga, M., ed.), MIT Press
5 Diamond, R. and Carey, S. (1986) Why faces are and are not special: an effect of expertise. *J. Exp. Psychol. Gen.* 115, 107–117
6 Grelotti, D. *et al.* (2002) Social interest and the development of cortical face specialization: what autism teaches us about face processing. *Dev. Psychobiol.* 40, 213–225
7 McCandliss, B.D. *et al.* (2003) The visual word form area: expertise for reading in the fusiform gyrus. *Trends Cogn. Sci.* 7, 293–299
8 Palmeri, T.J. and Gauthier, I. (2004) Visual object understanding. *Nat. Rev. Neurosci.* 5, 291–303
9 Goldstone, R.L. (1998) Perceptual learning. *Annu. Rev. Psychol.* 49, 585–612
10 Schyns, P.G. *et al.* (1998) The development of features in object concepts. *Behav. Brain Sci.* 21, 1–54
11 Barsalou, L.W. *et al.* (2003) Grounding conceptual knowledge in modality-specific systems. *Trends Cogn. Sci.* 7, 84–91
12 Jennings, C. and Aamodt, S. (2000) Computational approaches to brain function. *Nat. Neurosci.* 3 (Suppl), 1160
13 Kim, N.S. and Ahn, W-K. (2002) Clinical psychologists' theory-based representations of mental disorders predict their diagnostic reasoning and memory. *J. Exp. Psychol. Gen.* 131, 451–476
14 Tanaka, J.W. and Taylor, M. (1991) Object categories and expertise: is the basic level in the eye of the beholder? *Cogn. Psychol.* 23, 457–482
15 Johansen, M.K. and Palmeri, T.J. (2002) Are there representational shifts during category learning? *Cogn. Psychol.* 45, 482–553
16 Newell, A. and Rosenbloom, P.S. (1981) Mechanisms of skill acquisition and the law of practice. In Cognitive skills and their acquisition (Anderson, J.R., eds), pp. 1–55, Erlbaum
17 Heathcote, A. *et al.* (2000) The power law repealed: the case for an exponential law of practice. *Psychon. Bull. Rev.* 7, 185–207
18 Rickard, T.C. (2004) Strategy execution in cognitive skill learning: an item-level test of candidate models. *J. Exp. Psychol. Learn. Mem. Cogn.* 30, 65–82
19 Gauthier, I. and Tarr, M.J. (1997) Becoming a 'Greeble' expert: exploring mechanisms for face recognition. *Vis. Res.* 37, 1673–1681
20 Gauthier, I. *et al.* (1998) Training 'Greeble' experts: a framework for studying expert object recognition processes. *Vis. Res.* (Special issue 'Models of Recognition'). 38, 2401–2428
21 Jolicoeur, P. *et al.* (1984) Pictures and names: making the connection. *Cogn. Psychol.* 16, 243–275
22 Tanaka, J.W. (2001) The entry point of face recognition: evidence for face expertise. *J. Exp. Psychol. Gen.* 130, 534–543
23 Young, A.W. *et al.* (1987) Configural information in face perception. *Perception* 10, 747–759
24 Gauthier, I. *et al.* (2003) Perceptual interference supports a non-modular account of face processing. *Nat. Neurosci.* 6, 428–432
25 Goldstone, R.L. (2003) Learning to perceive while perceiving to learn. In *Perceptual Organization in Vision: Behavioral and Neural Perspectives* (Kimchi, R. *et al.*, eds), pp. 233–278, Erlbaum
26 Sowden, P.T. *et al.* (2000) Perceptual learning of the detection of features in X-ray images: a functional role for improvements in adults' visual sensitivity? *J. Exp. Psychol. Hum. Percept. Perform.* 26, 379–390
27 LeGrand, R. *et al.* (2001) Early visual experience and face processing. *Nature* 410, 890
28 Mondloch, C.J. *et al.* (2003) Developmental changes in face processing skills. *J. Exp. Child Psychol.* 86, 67–84
29 Tarr, M.J. and Gauthier, I. (2000) FFA: a flexible fusiform area for subordinate-level visual processing automatized by expertise. *Nat. Neurosci.* 3, 764–769
30 Palmeri, T.J. (1997) Exemplar similarity and the development of automaticity. *J. Exp. Psychol. Learn. Mem. Cogn.* 23, 324–354
31 Gauthier, I. *et al.* (1999) Can face recognition really be dissociated from object recognition? *J. Cogn. Neurosci.* 11, 349–370
32 Logothetis, N.K. and Pauls, J. (1995) Psychophysical and physiological evidence for viewer-centered object representations in the primate. *Cereb. Cortex* 5, 270–288

33 Gauthier, I. *et al.* (1999) Activation of the middle fusiform 'face area' increases with expertise in recognizing novel objects. *Nat. Neurosci.* 2, 568–573
34 Haxby, J.V. *et al.* (2001) Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293, 2425–2430
35 Nosofsky, R.M. (1986) Attention, similarity, and the identification-categorization relationship. *J. Exp. Psychol. Gen.* 115, 39–57
36 Riesenhuber, M. and Poggio, T. (1999) Hierarchical models of object recognition in cortex. *Nat. Neurosci.* 2, 1019–1025
37 Tarr, M.J. and Yi, D.C. (2003) Learning to see faces and objects. *Trends Cogn. Sci.* 7, 23–30
38 Ashby, F.G. (1992) Multidimensional models of perception and cognition, Lawrence Erlbaum Associates Inc
39 Hummel, J.E. and Biederman, I. (1992) Dynamic binding in a neural network for shape recognition. *Psychol. Rev.* 99, 480–517
40 Edelman, S. (1999) Representation and recognition in vision, MIT Press.
41 Poggio, T. and Edelman, S. (1990) A network that learns to recognize three-dimensional objects. *Nature* 343, 263–266
42 Ullman, S. *et al.* (2002) Visual features of intermediate complexity and their use in classification. *Nat. Neurosci.* 5, 682–687
43 Bülthoff, H.H. and Edelman, S. (1992) Psychophysical support for a two-dimensional view interpolation theory of object recognition. *Proc. Natl. Acad. Sci. U. S. A.* 89, 60–64
44 Tarr, M.J. *et al.* (1998) Three-dimensional object recognition is viewpoint dependent. *Nat. Neurosci.* 1, 275–277
45 Kruschke, J.K. (1992) ALCOVE: an exemplar-based connectionist model of category learning. *Psychol. Rev.* 99, 22–44
46 Nosofsky, R.M. and Palmeri, T.J. (1997) An exemplar-based random walk model of speeded classification. *Psychol. Rev.* 104, 266–300
47 Erickson, M.A. and Kruschke, J.K. (2002) Rule-based extrapolation in perceptual categorization. *Psychon. Bull. Rev.* 9, 160–168
48 Nosofsky, R.M. and Palmeri, T.J. (1998) A rule-plus-exception model for classifying objects in continuous-dimension spaces. *Psychon. Bull. Rev.* 5, 345–369
49 Ashby, F.G. *et al.* (1998) A formal neuropsychological theory of multiple systems in category learning. *Psychol. Rev.* 105, 442–481
50 Ashby, F.G. and Waldron, E.M. (1999) On the nature of implicit categorization. *Psychon. Bull. Rev.* 6, 363–378
51 Logan, G.D. (2002) An instance theory of attention and memory. *Psychol. Rev.* 109, 376–400
52 Logan, G.D. (1988) Toward an instance theory of automatization. *Psychol. Rev.* 95, 492–527
53 Rosseel, Y. (2002) Mixture models of categorization. *J. Math. Psychol.* 46, 178–210
54 Sigala, N. and Logothetis, N.K. (2002) Visual categorization shapes feature selectivity in the primate temporal cortex. *Nature* 415, 318–320
55 Nosofsky, R.M. (1998) Optimal performance and exemplar models of classification. In Rational models of cognition (Oaksford, M. and Chater, N., eds), Oxford University Press
56 Usher, M. and McClelland, J.L. (2001) On the time course of perceptual choice: the leaky competing accumulator model. *Psychol. Rev.* 108, 550–592
57 Lamberts, K. (2000) Information-accumulation theory of speeded categorization. *Psychol. Rev.* 107, 227–260
58 Nosofsky, R.M. and Alfonso-Reese, L.A. (1999) Effects of similarity and practice on speeded classification response times and accuracies: further tests of an exemplar-retrieval model. *Mem. Cogn.* 27, 78–93
59 Nosofsky, R.M. (1987) Attention and learning processes in the identification and categorization of integral stimuli. *J. Exp. Psychol. Learn. Mem. Cogn.* 13, 87–108
60 Maddox, W.T. (2002) Learning and attention in multidimensional identification and categorization: separating low-level perceptual processes and high-level decisional processes. *J. Exp. Psychol. Learn. Mem. Cogn.* 28, 99–115
61 Dosher, B.A. and Lu, Z.L. (1999) Mechanisms of perceptual learning. *Vis. Res.* 39, 3197–3221
62 Shepard, R.N. (1987) Toward a universal law of generalization for psychological science. *Science* 237, 1317–1323
63 Ennis, D.M. (1988) Toward a universal law of generalization. *Science* 242, 944
64 Livingston, K.R. *et al.* (1998) Categorical perception effects induced by category learning. *J. Exp. Psychol. Learn. Mem. Cogn.* 24, 732–753
65 Diehl, R.L. *et al.* (2004) Speech perception. *Annu. Rev. Psychol.* 55, 149–179
66 Goldstone, R.L. *et al.* (2001) Altering object representations through category learning. *Cognition* 78, 27–43
67 Nosofsky, R.M. *et al.* (1994) Comparing models of rule-based classification learning: a replication and extension of Shepard, Hovland, and Jenkins (1961). *Mem. Cogn.* 22, 352–369
68 Goldstone, R.L. and Steyvers, M. (2001) The sensitization and differentiation of dimensions during category learning. *J. Exp. Psychol. Gen.* 130, 116–139
69 Gauthier, I. and Palmeri, T.J. (2002) Visual neurons: categorization-based selectivity. *Curr. Biol.* 12, R282–R284
70 Op de Beeck, H. *et al.* (2003) The effect of category learning on the representation of shape: dimensions can be biased, but not differentiated. *J. Exp. Psychol. Gen.* 132, 491–511
71 Donnelly, N. and Davidoff, J. (1999) The mental representations of faces and houses: Issues concerning parts and wholes. *Vis. Cogn.* 6, 319–343
72 Farah, M.J. *et al.* (1998) What is 'special' about face perception? *Psychol. Rev.* 105, 482–498
73 Shiffrin, R.M. and Lightfoot, N. (1997) Perceptual learning of alphanumberic-like characters. In *Perceptual Learning: The Psychology of Learning and Motivation* (Vol. 36) (Goldstone, R.L., ed.), pp. 45–81, Academic Press
74 Goldstone, R.L. (2000) Unitization during category learning. *J. Exp. Psychol. Hum. Percept. Perform.* 26, 86–112
75 Wenger, M.J. and Ingvalson, E.M. (2003) Preserving informational separability and violating decisional separability in facial perception and recognition. *J. Exp. Psychol. Learn. Mem. Cogn.* 29, 1106–1118
76 Keri, S. (2003) The cognitive neuroscience of category learning. *Brain Res. Brain Res. Rev.* 43, 85–109
77 Smith, P.L. and Ratcliff, R. (2004) Psychology and neurobiology of simple decisions. *Trends Neurosci.* 27, 161–168
78 Medin, D.L. *et al.* (2002) Categorization and reasoning in relation to culture and expertise. In *The Psychology of Learning and Motivation* (Vol. 41) (Ross, B.H., ed.), pp. 1–41, Academic Press
79 Gauthier, I. *et al.* (2000) The fusiform 'face area' is part of a network that processes faces at the individual level. *J. Cogn. Neurosci.* 12, 495–504
80 Grill-Spector, K. and Kanwisher, N. Visual recognition: as soon as you know it is there, you know what it is. *Psychol. Sci.* (in press)