BEYOND FACES AND MODULARITY

(for Trends in Cognitive Science)

Cindy Bukach\textsuperscript{1,3}
Isabel Gauthier\textsuperscript{1,3}
Michael J. Tarr\textsuperscript{2,3}

\textsuperscript{1}Department of Psychology, Center for Integrative and Cognitive Neuroscience, Vanderbilt Vision Research Center, Vanderbilt University, Nashville, TN

\textsuperscript{2}Department of Cognitive and Linguistic Sciences, Brown University, Providence, RI

\textsuperscript{3}Perceptual Expertise Network

August 15, 2005

Address Correspondences to:
Cindy Bukach
Department of Psychology
301 Wilson Hall
Vanderbilt University
Nashville, TN 37240
tel: 615-322-5265
fax: 615-322-4706
cindy.bukach@vanderbilt.edu
The cognitive neuroscience of visual cognition often focuses on category differences, contrasting the behavioral and neural signatures associated with different object categories such as faces, objects, letters or scenes. In contrast, our approach emphasizes the cognitive and neural mechanisms recruited by observers with substantial experience – perceptual expertise – performing a given task with objects from a particular category. Beyond offering an alternative to the domain-specific account of face specialization, the perceptual expertise framework is a systematic approach to the general phenomenon of category specialization as it arises from interactions between experience and biases, providing a window into the functional plasticity of the mind and brain.

Our study of perceptual expertise has emerged from the significant body of research on face-specificity that reveal both behavioral (e.g., holistic and relational effects) and neural (e.g., the fusiform face area or FFA found with fMRI, the N170 ERP component) markers important for the skilled individuation of faces[1-5]. The study of face recognition continues to provide vital clues as to how higher-level vision is organized. In this spirit, we conceptualize face processing as one instance of perceptual expertise, extending such research beyond the face domain. Studies of expertise have also extended our understanding of face processing by demonstrating that putatively face-specific markers arise in the skilled individuation of other homogeneous categories, such as birds, cars, dogs, and even novel objects such as Greebles (Figure 1)[6-10]. For example, the right FFA of car and bird experts is more active when viewing objects within their domain of expertise as compared to viewing objects from another homogeneous object category[7]. Importantly, individual differences in FFA activity are significantly correlated with independent behavioral tests of expertise, a result obtained in both blocked and event-related
designs that control for attentional confounds[6, 11]. More abstract individuation tasks also give rise to behavioral and neural effects typically associated with face processing. For example, fingerprint experts, but not novices, process fingerprints holistically and show a delayed N170 potential when fingerprints are inverted[12], a result obtained with both faces[13] and Greebles[14].

Such findings have obvious relevance regarding the domain specificity of mechanisms involved in face processing and also address more fundamental questions across domains, that is, the principles of functional and cortical plasticity underlying expertise acquisition. As such, our goal is not to demonstrate that all instances of expertise are computationally and neurally isomorphic, but rather to identify the cognitive and neural mechanisms recruited by expertise acquisition within visual domains as varied as face, object and letter recognition, as well as in non-visual domains such as bird song.

**Acquisition of Expertise**

The expertise approach provides a tool for studying behavioral and neural changes as learning occurs in a laboratory setting, where both training procedures and stimulus properties can be controlled and manipulated. A typical model for expertise training involves discrimination training at both categorical and individual levels, continuing until response times are equivalent for the two conditions[15]. One surprising finding has been the striking behavioral and neural changes we observe over a relatively short time period. That is, the training our subjects receive in the laboratory is far shorter than the 10+ years once thought to be necessary for the acquisition of expertise[10]. For example, studies using Greebles have revealed both neural and behavioral changes after only 10 *hours* of laboratory training. We have identified at
least two brain regions that become selectively more active to Greebles as Greeble expertise is acquired: the FFA, and the “occipital face area” (OFA)[15]. Three behavioral markers of expertise have also been obtained with the onset of expertise: a subordinate-level shift, holistic processing, and relational processing[15-17] (Box 1). We are also trying to uncover the relationship between such behavioral markers and observed neural changes. One study found changes in FFA activity correlated with behavioral measures of holistic processing, but not relational processing, suggesting that the two may be dissociable[17] – supported by the recent finding that FFA activity is not correlated with relational face processing[18]. The role of the OFA in expertise acquisition is less well understood, but may be involved in the early perception of individual object features[19].

Generalization Within Object Domains

Another issue our research addresses is the generalization of expertise to new exemplars. It is the generalization to new exemplars within a trained category that distinguishes expertise from stimulus-specific perceptual learning[20]. Thus an important characteristic of our training studies is that the behavioral and neural effects of expertise training are assessed using untrained (unfamiliar) exemplars.

Research on perceptual skill acquisition often conceptualizes categories according to broad definitions (e.g., faces, dogs, cars, butterflies), but it is not the case that perceptual skills necessarily generalize to all exemplars within such broad domains[10]. Even within the domain of faces, limits of generalization are evident. For example, we are much better at remembering faces of our own race than faces of other races[21] and the FFA and other face-selective regions show higher activation to own-race than to other-race faces[22]. Such effects can be explained
by differences in experience that lead to different processing styles for own- versus other-race faces. For example, in the part-whole task (Box 1), Caucasians who had little prior exposure to Asian faces showed greater holistic processing for Caucasian faces than for Asian faces[23] and, more importantly, the level of holistic processing for other-race faces was related to experience: Asian subjects who had significant prior exposure to both Asian and Caucasian faces showed equivalent holistic processing for Asian and Caucasian faces. Thus, even within the highly expert domain of faces, experience affects how different exemplars are processed and, consequently, the generalization of perceptual expertise.

Expertise studies with non-face objects have identified two factors that contribute to whether expertise generalizes across exemplars within a given domain:

i. The type of experience (individuation vs. mere exposure). A subject’s ability to generalize may depend on how often particular exemplars are individuated at a more specific level; mere exposure may not be sufficient[24]. For instance, bird novices were trained on owls and wading birds. One category was trained at the subordinate species level (“great grey owl vs. eastern screech owl”), and the other at the basic family level (owl vs. wading bird). Following training, subjects showed good discrimination skills for new exemplars and species for the category trained at the species-level, but not to the category trained at the family level.

ii. Geometric and surface similarity. Greeble training studies have highlighted the importance of geometric and surface similarity for generalization. Expertise generalized to new Greebles that were geometrically similar to those used during training, but did not generalize to new Greebles that were distinguishable only by information that was not recruited during training tasks[15]. These studies not only reveal important principles of visual learning, but may
help to account for some of the conflicting evidence regarding the relationship between expertise for faces and non-face objects (Box 2).

**Impaired Processing**

The expertise approach has also helped us to better understand the underlying cognitive impairments in category-specific deficits such as prosopagnosia. For example, a study with two prosopagnosic subjects performing a task that manipulated discrimination difficulty *within* object classes revealed that these individuals had deficits in fine-level discrimination for both faces and non-face objects[8], thereby implicating a domain-general impairment. Even more revealing is the use of expertise training in subjects with visual recognition deficits, a technique that can lead to novel predictions about unimpaired object recognition. A recent study found that expert processing is spatially limited in some types of prosopagnosia: whereas fine-level discrimination of relational and featural information was severely impaired in the eye region, processing of fine-grained details was preserved for the lower region of the face. Although this prosopagnosic subject was eventually able to reach criterion following extensive Greeble training, identification of both faces and Greebles was based primarily on a single feature (the mouth for faces, and the upper appendage for Greebles)[25]. This spatially graded loss of expertise raises the possibility that expertise mechanisms may also be acquired in a spatially graded manner such that intermediate stages of training may give rise to expertise for a restricted spatial area. Consistent with this hypothesis, although Greeble training studies with longer training protocols show holistic processing across the entire Greeble[16], a study with a shorter training protocol resulted in holistic processing for only the upper half of each Greeble [17].
Of course, face recognition deficits may arise due to a number of causes, and may not necessarily involve perceptual or expert mechanisms[26]. For example, DD, an individual diagnosed with Autism, shows no evidence of face expertise and no face selectivity in the fusiform gyrus, yet exhibits robust fusiform selectivity in response to cartoon characters for which he is an expert[27]. This pattern suggests that DD’s face deficit is not due to impaired expertise mechanisms *per se*, but to mechanisms that modulate interest to faces. This finding raises the intriguing possibility that individuals with Autism may benefit from expertise training with faces. That is, a training protocol could supply the external motivation to support the sort of experiences necessary to develop expertise with faces[28]. Consistent with this conjecture, our finding of expertise effects in laboratory-trained experts with artificial objects suggests that, although expertise typically arises from intrinsic interest in a domain (e.g., a hobby), such motivational factors may not be necessary for learning to occur. As such, our approach moves beyond the domain-specific debate by identifying specific cognitive and neural processes; in this example, those that underlie category-specific impairments. This and related approaches may provide useful diagnostic and rehabilitation tools (e.g., rehabilitating face skills in cases where expertise mechanisms are intact).

**Varieties of Expertise**

In much of our research we have emphasized the fact that visual expertise across many different categories recruits similar cognitive and neural mechanisms. However, it is not the case that all types of expertise rely (or should rely) on identical mechanisms. Expertise with stimuli that vary radically from the geometry and functional goals of homogeneous object individuation may not engage the FFA, but may recruit other, functionally appropriate, regions. For example, words, unlike faces, do not share a first order configuration. Rather, the meaning of words
depends upon the identity and the serial order of a limited number of letters. Moreover, word reading must generalize across changes in capitalization and font. Thus it is not surprising that expert word reading is not associated with the FFA, but with the visual word form area (VWFA), a region in the left fusiform gyrus[29]. Recent studies have also dissociated expert word reading from expert single letter identification: Whereas word reading is associated with selectivity in the VWFA, single letter identification is associated with selective activity in a different part of the left fusiform[30, 31]. At the same time, expert letter recognition is associated with enhanced N170 responses bilaterally[32], suggesting that this ERP component is indicative of expert processes for different categories (e.g., faces, objects, fingerprints, letters) that may be localized in different brain areas. Still another type of visual expertise relies on local shape contrasts rather than configural information[33]. It seems unlikely that such expertise would recruit FFA any more than basic-level object recognition.

To this point, we have focused on visual expertise. A similar approach can be used to study cortical plasticity in other domains. Chess expertise, sometimes thought of as a cognitive skill, may actually have a significant visual component[34]. Indeed, the rapid recognition of the complex spatial relations between pieces may be critical in playing chess at the expert level. As such, it is not surprising that a recent study found that chess experts, but not novices, show category selectivity for valid chess boards (but not invalid chess boards) in the region of the right fusiform gyrus typically associated with face processing[35]. A different sort of expertise may be found for “birders”, some of whom are able to identify individual bird species through both vision and audition (song). Consistent with brain organization by modality, bird song elicits selective activity in auditory cortex that is modulated by expertise (Brodmann Area 22). Interestingly, birdsong also elicits expertise-related activation in both lower- and higher-level
visual areas as well, including a region in the right fusiform gyrus[36]. This fusiform activation is adjacent to, but not overlapping with the functionally defined FFA. Similar auditory-related activation in regions adjacent to the FFA has been found for the auditory recognition of familiar voices[37]. Such expertise-related cross-modal activation may arise due to multimodal representations for specific objects (e.g., a given person or bird species), to visual imagery elicited by auditory stimulation, or to domain-general expertise mechanisms within the fusiform gyrus. Conversely, images of faces and objects have recently been found to elicit distinct responses in the auditory cortex of the macaque, not unlike the category-selective responses typically found in visual areas such as the superior temporal sulcus[38]. Finally, behavioral markers associated with expert object recognition are also obtained when subjects are trained to identify 2-D and 3-D objects by touch. For example, experts but not novices show better tactile recognition of upright than inverted patterns, and of object wholes as opposed to object parts[39]. It is unknown exactly which neural substrates are recruited by tactile expertise, as well as whether these overlap, at least in part, with the substrates recruited by expertise in the visual domain. Interestingly, research across modalities can be used to examine whether expertise in a given domain in one modality transfers to that domain in another modality[40, 41].

**Defining the Relationship between various types of expertise**

One recurring issue is the functional and neuroanatomical relationship between separable domains of expertise. For example, do faces, birds, cars, Greebles, etc. rely on shared cognitive mechanisms and/or shared neural substrates? Several recent studies have addressed this question by examining whether there is spatial overlap between the category-selective regions in the fusiform gyrus for faces and other domains of expertise. Two different fMRI studies (one using
butterflies, another using cars) found that the overlap in activation for faces and for the expert domain was smaller than the overlap for faces across different blocks or tasks[18, 42]. However, because we understand so little about neural coding, it is difficult to know how to interpret these results. First, there was some overlap between domains and it is unclear what degree of overlap (or lack thereof) signifies independent mechanisms. Indeed, there is good evidence that neural codes for object categories are somewhat distributed, yet still overlapping[43, 44]. Second, many visual areas, for example V1, exhibit spatially selective coding for different stimulus properties, yet functional equivalence across a much wider brain region[45]. Thus, even completely non-overlapping regions may not necessarily signify functional independence.

An alternative to measuring spatial overlap is examining functional overlap through dual task paradigms. If the perception of two objects share overlapping cognitive and neural mechanisms (as opposed to neural locations), then a task that requires simultaneous processing of both objects should reduce the availability of the shared process applied to either (providing that the mechanisms in question have limited capacities). The efficacy of this approach has been demonstrated in an ERP study that found that the N170 in response to faces was significantly reduced when observers were simultaneously required to process a second face, but was not affected when the second face was scrambled[13]. Simultaneous processing of a Greeble with a face also reduces the N170 response to faces for Greeble experts[46]. A variant of this technique has been used to examine whether a specific mechanism recruited by visual expertise, holistic processing (Box 1), shares a common neural substrate in the expert processing of faces and cars[47]. Car experts showed lower holistic processing for faces and reduced N170 amplitudes to faces when they had to process cars and faces simultaneously (as well as greater holistic processing for cars and a higher N170 amplitude for cars as compared to novices). The degree of
interference between faces and cars within an individual was positively correlated with an
independent measure of car expertise. Thus, the neural processes used by car experts for holistic
processing of cars are not functionally independent from those that support holistic face
processing.

Although the majority of work to date has focused on the issue of domain specificity for
faces, the same techniques may be used to identify functional similarities and differences
between any two domains of expertise, both across categories and modalities. For example,
interference paradigms allow us to assess whether any two skills depend on different functional
modules. Such an approach aids in establishing whether specific cognitive functions are shared
or independent, as well as allowing us to better understand how related cortical functions are
spatially arranged in the brain.

Towards a Taxonomy of Expertise

The combination of evidence from studies of face-selectivity, of extant expertise for non-
face objects, and of expertise acquired in the laboratory, provides a basis for a taxonomy of the
cognitive and neural mechanisms engaged by different forms of perceptual expertise. Although
much work remains, recent work suggests that this taxonomy will likely reflect the physical and
conceptual properties of the stimulus class (e.g., object geometry, modality, and functional
knowledge); task and/or experiential factors (e.g., level of individuation); as well as inherent
biases due to principles of neural organization (e.g., modality, eccentricity, hierarchy[48, 49]).
The development of this taxonomy will continue to necessitate the comparison of cognitive and
neural mechanisms for different categories and sensory domains. Of course such research
remains relevant to the question of domain specificity, however, we believe that the value of the
expertise framework moves beyond demonstrations of isomorphism. Our approach is useful in
its own right as a tool for exploring the cognitive and cortical mechanisms underlying expertise acquisition, and as a means to study the computational properties and cortical plasticity of the primate brain.
Box 1. Behavioral Markers of Expertise

Subordinate-level shift: Novices identify objects at the basic level more efficiently than at a more specific, subordinate level[50], but expertise can lead to faster responses at the subordinate level[51]. This entry-level shift manifests itself as reduced response times to subordinate-level judgments as expertise is acquired (Figure 1).

Holistic Processing: Experts encode information from a wider spatial extent as compared to novices. As such, information across the entire object is likely to affect their behavior and they are less able to selectively attend to single parts[17]. Methods for measuring holistic processing include the Part-Whole (Figure 2) and the Composite (Figure 3) tasks.

Relational Processing: Experts are more likely to encode information about the spatial relations between parts of an object[17]. Relational processing can be indexed by sensitivity to spatial changes between the parts of an object using methods such as the Isolated Change (Figure 4) or Composite (Figure 5) tasks.
Box 2. When Experts Perform Like Novices.

Despite numerous studies demonstrating that expertise with various homogeneous categories modulates FFA response[6, 8, 9], the validity of the expertise hypothesis continues to be controversial. For example, some studies fail to find expertise effects in the FFA for certain non-face object categories[18, 42, 52]. However, such studies surprisingly fail to show any expertise-related neural activity, even outside face-selective areas, a result that can be difficult to interpret. On the other hand, this raises the interesting possibility that experts sometimes perform like novices within their domain of expertise. Although it has been hypothesized that experts engage expertise mechanisms automatically, recent evidence suggests that task demands can affect the degree to which these mechanisms are engaged. For example, the FFA is more active during an old/new recognition task than during a passive viewing task, for both expert and non-expert categories (including objects, lepidoptera, and faces)[42]. Recruitment of expert mechanisms is also dependent on stimulus properties: Expert effects do not fully generalize across different races of faces,[23] dog species,[10] bird species[24, 53] or Greeble sets with different geometric properties[15]. Such limits to the generalization of expertise suggest an explanation for why some studies have failed to replicate expertise-related activation in the FFA of real-world experts. In particular, a study of Lepidoptera experts[42] used Lepidoptera species that were unfamiliar to the experts and a study with modern car experts[18] used, in the fMRI tasks, mostly antique cars. Importantly, a better understanding of the conditions under which expert mechanisms are engaged or not will also enhance our understanding of object recognition more generally.
Box 3. Questions for future research

What is the role of decisional processes in expertise?

What is the role of attention in expertise?

How does feature saliency change with expertise?

Can we create feature-based expertise, and how will it differ from holistic expertise?

What role do semantic features play in perceptual expertise?

How does expertise for different modalities (visual, auditory, tactile) differ? In what ways are they the same?

How does expertise affect working memory capacity?

Can we dissociate FFA activity related to detection from FFA activity related to identification?

How do interactions between the amygdala and the FFA influence the acquisition of expertise?

Are there special populations that can benefit from face training?
Figure Captions

*Figure 1.* Examples of symmetrical and asymmetrical Greebles. Greebles are homogeneous novel objects that can be individuated by the shape of their appendages, and grouped into different families according to body type. The top row shows one sample symmetrical Greeble from each of five families. The bottom row shows one sample asymmetrical Greeble from each of five families. Greebles can be further classified by gender according to the orientation of their appendages (all Greebles shown here are female).

Box 1 Figure Captions

*Figure 1.* Hypothetical data showing subordinate-level shift.

*Figure 2.* Part-Whole Task. Holistic processing is indicated by a greater advantage for part recognition when the test stimuli are wholes, than when the test stimuli are isolated parts.

*Figure 3.* Composite Task. Holistic processing is indicated by greater interference from the irrelevant half in a parts-matching task (performance is poorer when responses to the bottom and top of stimuli are incongruent than when they are congruent).

*Figure 4.* Isolated Change Task. Sensitivity to *local spatial information* can be measured by a matching task in which the spatial relations between parts have been manipulated.

*Figure 5.* Relational Processing in the Composite Task. Sensitivity to *global spatial information* can be measured by a release from interference when object parts are misaligned.
Figure 1
Box 1. Behavioral Markers of Expertise

**Subordinate-level shift:** Novices identify objects at the basic level more efficiently than at a more specific, subordinate level[50] but expertise can lead to faster responses at the subordinate level[51]. This entry-level shift manifests itself as reduced response times to subordinate-level judgments as expertise is acquired (Figure 1).

**Holistic Processing:** Experts encode information from a wider spatial extent as compared to novices. As such, information across the entire object is likely to effect their behavior and they are less able to selectively attend to single parts[17]. Methods for measuring holistic processing include the Part-Whole (Figure 2) and the Composite (Figure 3) tasks.

**Relational Processing:** Experts are more likely to encode information about the spatial relations between parts of an object[17]. Relational processing can be indexed by the sensitivity to spatial changes between the parts of an object using such methods as the Isolated Change (Figure 4) and the Composite (Figure 5) tasks.

---

*Figure 2.* Part-Whole Task. Holistic processing is indicated by a greater advantage for part recognition when the test stimuli are wholes, than when the test stimuli are isolated parts.

*Figure 3.* Composite Task. Holistic processing is indicated by greater interference from the irrelevant half in a parts-matching task (performance is poorer when responses to the bottom and top of stimuli are incongruent than when they are congruent).

*Figure 4.* Isolated Change Task. Sensitivity to *local spatial information* can be measured by a matching task in which the spatial relations between parts have been manipulated.

*Figure 5.* Relational Processing in the Composite Task. Sensitivity to *global spatial information* can be measured by a release from interference when object parts are misaligned.